

PCC - Codebook V3.2.01

The PCC project team

January 9, 2018

Contents

List of Tables	3
1 Project Description	4
1.1 Aim of this Document	4
1.2 Short Project Summary	4
1.3 Mid-term Roadmap of the Project	4
2 Data Structure, Sources & Sampling	4
2.1 Data Structure	4
2.2 Data Sources	5
2.3 Sample	5
3 How to Read this Codebook	5
3.1 Labels, Abbreviations, Etc.	5
3.2 Collection Effort Per Variable	5
4 Politician Level Data Frame (POLI)	7
5 Parliamentary Episode Data Frame (PARE)	9
6 Resume Entries Data Frame (RESE)	11
6.1 Political functions	14
6.1.1 Principles for Classification of Political Functions	14
6.1.2 Primary Political Function Coding	14
6.1.3 Secondary Political Function Codes	27
6.1.4 Secondary Political Function Coding: Bi-Cameral System (Lower House, Upper House) Applies to: X_LE:	30
6.1.5 Political functions coding: some final specifications	30
7 Parliament Data Frame (PARL)	31
8 Membership Episode Data Frame (MEME)	32
9 Party Data Frame (PART)	33
10 Faction Data Frame (FACT)	36
11 Committee Data Frame (COMM)	37
12 Electoral List Data Frame (ELLI)	40
13 Electoral List Entry Data Frame (ELEN)	42
14 Electoral District Data Frame (ELDI)	44
15 Organisations Module(ORGS)	45
16 The Quota Module (QUOT)	50
17 Appendices	51
17.1 Alternative IDs: ALID	51
17.2 Regional Abbreviations Used	51
17.2.1 Regional Abbreviations Used for CH	51
17.2.2 Regional Abbreviations Used for DE	52
17.2.3 Regional Abbreviations Used for NL	52

List of Tables

1	Summary of Politician Level Variables (POLI)	7
2	Summary of the Parliamentary Episode Level Variables (PARE)	9
3	Summary of the Resume Entry Variables (RESE)	11
4	Political functions level 3 coding specifications	21
5	Example of Three-tier Coding	21
6	Example of Secondary Coding	22
7	Political functions level 4 coding specifications: policy area codes Note: The letter S in front of the numerical code specifies newly created codes based on the U.S: SIC classification.	26
8	Level 5: Occupational Position Codes	26
9	Elected or Non-Elected Political Functions - DNC	28
10	Specification of political function secondary code: legislative steering body codes	28
11	Youth or Regular Party Political Functions	28
12	Shadow or Regular Political Functions - DNC	29
13	Important Party Committees	29
14	Specification of political function secondary code: interest group codes	30
15	Summary of Parliament Level Variables (PARL)	31
16	Summary of Membership Episode Variables (MEME)	32
17	Summary of Party Variables (PART)	34
18	Summary of Faction Variables (FACT)	36
19	Summary of Faction Variables (COMM)	38
20	Summary of the Electoral List Level Variables (ELLI)	40
21	Summary of the Electoral List Entry Variables (ELEN)	42
22	Summary of District Level Variables (ELDI)	44
23	Summary of Organisation Level Variables (ORGS)	46
24	Summary of Quota Variables (QUOT)	50
25	Alternative IDs	51

1 Project Description

1.1 Aim of this Document

The primary purpose of this codebook is to provide detailed information on the data structure and variables of the parliamentary careers in comparison data. This codebook's secondary - much more ambitious, long-term and as such necessarily more controversial - purpose is to establish a new¹ data-standard. As such, we did aim the design a data-structure and coding-standard that can be used to study political careers in any context (country, level, system...). Any input that informs our success in this regard is very welcome!

1.2 Short Project Summary

The research project "Parliamentary Careers in Comparison" (PCC) aims to investigate political careers and activities of parliamentary candidates and parliamentarians in Switzerland, Germany and the Netherlands² since the Second World War. Based on an extensive collection of detailed individual-level data, we investigate biographical and behavioural dynamics to obtain a full and dynamic picture of parliamentary careers. The research interests covered by the project cover three broad areas. The first set of questions revolves around career paths on different levels, how they typically develop, and how these patterns can be explained. With the second set of questions we focus on the impact of political institutions on political careers while a third strand highlights the consequences of the different career paths on parliamentary behaviour and future post-parliamentary careers.

The PCC data-set contains detailed information on the political careers of parliamentarians and parliamentary candidates. Examples of variables include socio-demographic information of national candidates and parliamentarians of the three countries since the second world war, the candidate list positions, political functions (if elected) of national parliamentarians. Future versions of this data will also include data on regional parliamentarians and expand the range of included variables. See parliamentarycareersincomparison.org for more information.

1.3 Mid-term Roadmap of the Project

Data will be collected in several waves over a three year period (2016-2019). The codebook below summarizes the data-collections efforts in the first wave. Future waves will likely include the collection of both *similar data on other politicians* (e.g. candidates for the national parliament, regional parliamentarians) and *different data for the same politicians* (e.g. additional variables in parliamentary behavior). Please drop us an email if you would like to see specific variables included. We can't promise anything but are always happy to consider!

2 Data Structure, Sources & Sampling

2.1 Data Structure

The PCC data-structure is organized into four levels and nine data-frames. Examples of these data-frames can be found [here](#) (.xlsx download). The logic behind this data-structure is to collect static and time-varying variables as much is possible in an 'intuitive' way, sticking close to the structures that researchers are used to and the format that different kinds of data (e.g. biographical profiles, electoral lists) are typically stored in. In practice this implies that each data frame only contains information varying on the same level (e.g. the individual or the electoral list). We use a system of identifiers to allow data-points to be connected up together. Figure 1 summarizes the data-frames and the relations between them.

- Politician level data
 - [Stable characteristics](#) - (see page 4) - POLI
 - [Parliamentary episodes](#) - (see page 5) - PARE
 - [Resume entries](#) - (see page 6) - RESE
 - [Membership episodes](#) - (see page 8) - MEME
- [Parties](#) (see page 9) - PART
- [Parliaments](#) (see page 7)- PARL

¹See Turner-Zwinkels 2016 for a specification of the requirements of such a new standard.

²Most of the Dutch data has already been collected as part of the PhD thesis of PCC project member Tomas Turner-Zwinkels.

- **Factions** (see page 10)- FACT
- **Committees** (see page 11)- COMM
- Electoral list data
 - **Electoral districts** (see page 14) - ELDI
 - **Electoral lists** (see page 12) - ELLI
 - **Electoral list entries** (see page 13) - ELEN

2.2 Data Sources

Data was collected by a combination of different **sources**. A main source for the CH and DE parliamentary career data are yearbooks. A main source in the Netherlands was the website of the Dutch parliamentary documentation centre. Other data stem from different statistical offices or are based on newspaper articles or personal websites.

2.3 Sample

Data can be regarded a **complete (population) sample of national parliamentarians since 1946 of all three countries (CH,DE³,NL)** (wave one data collection). Both parliamentarians in the lower-house (i.e. Bundestag, Nationalrat, Tweede Kamer) and the upper-house (i.e. Staenderat, Eerste Kamer) are included. Future waves will likely include candidates for national elections in Germany and Switzerland (and maybe the Netherlands) since 1946, and of regional parliamentarians of all three countries (wave two). The data will likely also include data on the candidates for elections to regional parliaments (e.g. German Bundeslaender or Swiss Cantons).

3 How to Read this Codebook

3.1 Labels, Abbreviations, Etc.

- Brackets (‘[’ and ‘]’) in variable names signify variable names. [country]_ for example means that there will be as many of this variable (set) in the data-frame as there are countries in the data.
- ‘DNC’ means: Do Not Collect, these are specifications that are kept in codebook for reference purposes who values we do not intend to collect (to this level of detail).
- if a variable name contains an (by underscores) indexed element on the lists which is not used ‘NA’ will be used a filler.

3.2 Collection Effort Per Variable

Not all variables are collected with the same effort. Our effort to collect different variables varies from the highest level (‘Collect and follow-up until COMPLETE’) to the lowest level (‘Collect When Easy’). These are the matching abbreviations:

- COMP - Collect and follow-up until complete
- CWA - Collect when available
- CWE - Collect when easy

³Up until 1990 only data on West-Germany is included. For obvious reasons.

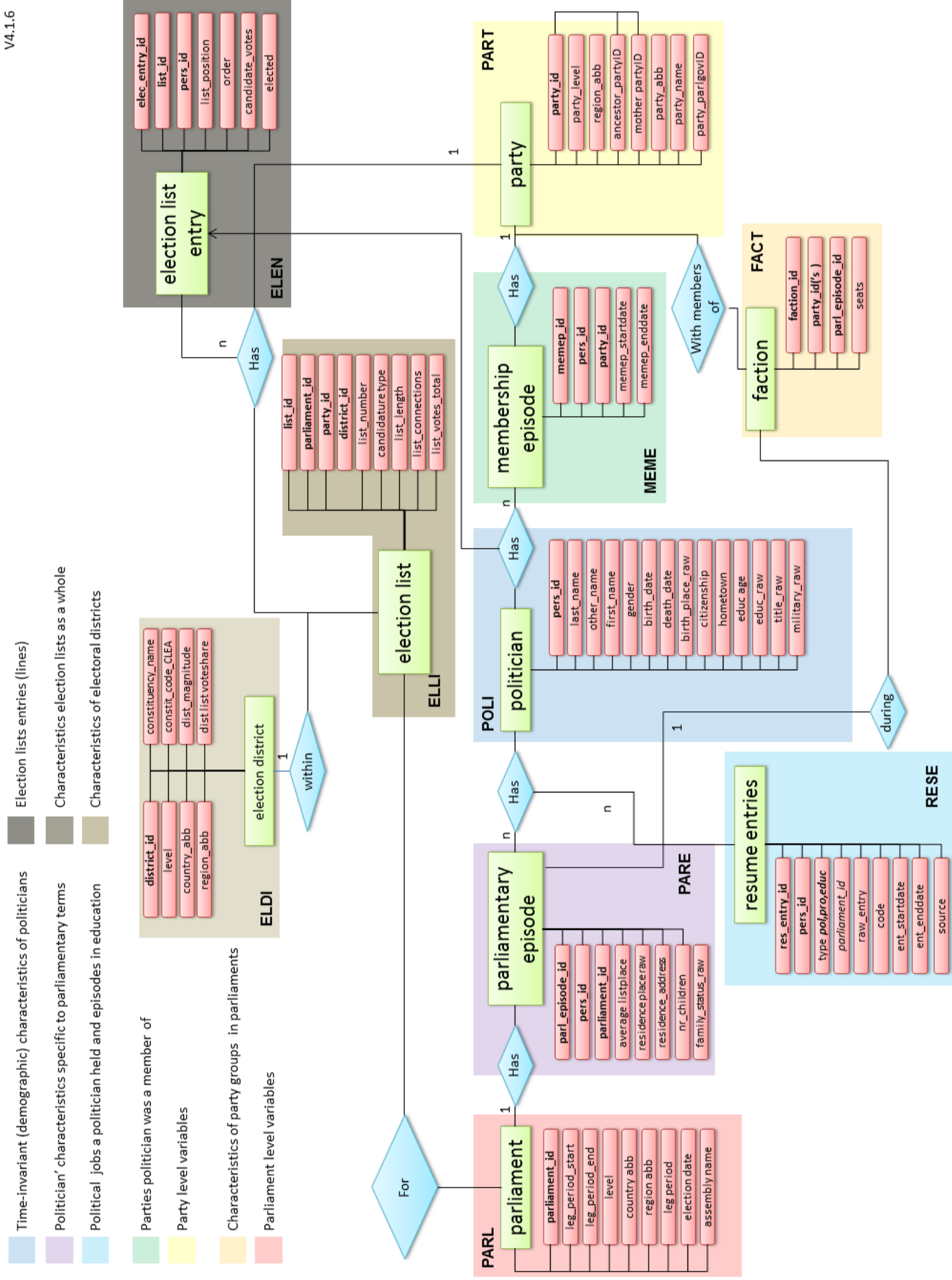


Figure 1: Entity Relationship Diagram of the data

4 Politician Level Data Frame (POLI)

Politician level variables are all statiC variables on the level of individual politicians. Example data can be found [here](#) (.xlsx download).

Table 1: Summary of Politician Level Variables (POLI)

NAME	TYPE	VALUES/EXAMPLE ⁴	SHORT DESCRIPTION	EFFORT
<code>pers_id</code> ⁶	PriID	[country] -[first_last_name] -[first_first_name] -[year_of_birth] CH_Abate _Fabio.1966 DE_Mueller _Johann.1956_dec	parliamentarians individual identification code: a combination of <code>country_abb</code> , the <code>last_name</code> , the “first_name” and the <code>birth_date</code>	COMP
<code>id_[country]_parl</code> ⁷	ID	2739 = Gobbi, Norman	parliamentsdienst CH identification number: identification number, as used by the CH Paliamentsdienst	CWA
<code>last_name</code>	String	Mueller deCourten Mutlu-Blum	parliamentarian’ last name: contains the last name. With ‘von’ e.t.c. connected (space removed), double last names hyphenated and special characters replaced.	COMP
<code>first_name</code>	String	“Jules-Henri” “Elisabeth” “Johann-Ulrich-Werner” “Heinrich-E.”	parliamentarian’ first name: contains the first name. With double first names hyphenated and special characters replaced.	COMP
<code>other_name</code>	String	Hans Muller	parliamentarian’ alternative name or alias: contains an (array of) - first or last - name(s) or alias(ses). For matching purposes.	CWA
<code>gender</code>	Binary string	m = male f = female	parliamentarian’ gender	COMP
<code>birth_date</code>	Date	16apr1897 29may1930	parliamentarian’ birth date	COMP
<code>death_date</code>	Date	16apr1897 29may1930	date of death of a parliamentarian 99 if still alive and NA when unknown	CWA
<code>birth_place_raw</code>	String	“Berlin” “Zuerich” “Baden-Wuerttemberg”	place of birth: place of birth of the parliamentarian	CWA
<code>citizenship</code>	String	DE DE; SK CH	parliamentarians citizenship(s): Array of - citizenship(s) as country abbreviation(s)	CWA

⁶To ensure compatibility of datasets and make usage and merging of our dataset easier we will add several id variables. The first one (`pers_id`) will be constructed by us, consists of the country, last name, first name, and birth year (and possibly a running number), and is unique to parliamentarians across countries and levels. This “information-based” ID format allows consistent IDS to be constructed across a variety of sources while minimizing the need to look-up IDS in the already collected data.

⁷We will add all important id variables (often numbers) that are used by important data sets and/or by the source data set. These additional id variables are mostly unique to a country and/or level and are named by the following principle: first “id”, then a country abbreviation (see variable “country_abb”) and then a name or names that indicate the usage or source, all added by underscores.

Table 1: Summary of Politician Level Data-frame (Continued)

NAME	TYPE	VALUES/EXAMPLE ⁵	SHORT DESCRIPTION	EFFORT
<code>citizenship_canton</code>	String	“ZH”	Swiss canton of citizenship: Switzerland-specific variable. Array of - citizenship(s) as canton abbreviation(s). Contains all citizenships over the entire lifetime.	CWA
<code>hometown_raw</code>	String	“Fehraltorf”	Hometown of a person: Switzerland-specific string variable. Name (and description) of place/city/municipality that is the candidates/parliamentarians home town.	CWA
<code>educ_age</code>	Integer	0 - 100	age at completion of education: ratio variable indicating how old a politician was when he/she finished her last degree	CWA
<code>educ_raw</code>	String	School ‘The Bear’ 1932	education: raw text string field with all available educational information	COMP
<code>title_raw</code>	String	Prof. Dr.	academic title: Academic title(s), if any. ⁸	COMP
<code>military_raw</code>	String	Major Appointé	military rank: raw string information about the military career	CWA

pers_id **parliamentarians individual identification code:** Identification code, individual to each parliamentarian ACROSS levels. The code consists of the country abbreviation (as specified in variable `country_abb`), the last name (as specified in variable `last_name`). The first name (as specified in `first_name`) and the year of birth (as specified by the last four digits of variable `birth_date`) connected by underscores. In case this id is not unique to one person, the month of birth (as specified in the middle of the `birth_date` variable string) is appended by underscore. If the month of birth is not known or does still not uniquely identify the person, the party abbreviation (as specified by the variable `party_abb`) is appended (additionally to the month of birth) by underscore. Alternative IDs and how they correspond to the main ID can be found in the appendix.

last_name & first_name **For the cleaning of names** the following rules apply:

- Names always start with a capital letter
- an umlaut (ö,ä,ü) is written as “oe”, “ae”, “ue”
- a “ß” will be written as “ss”
- all accents or similar (é, ã, ê, š) are left out, instead just the basic letter is written (e, a, e, s).
- if names contains prepositions like ‘Von ‘ or ‘ Van der’ then the space between these prepositions (name-particle) and the actual lastname are removed and capitals are replaced. So, ‘Von Liebig’ is spelled as ‘vonLiebig’ and ‘Van Der Maden’ is spelled as ‘vanderMaden’.
- if the names contain multiple parts (e.g. first_name Jean Luc) then the spaced are replaced with hypens (e.g. Jean-Luc)
- Additional non-name particles such as junior are included after underscores (Carl_jun.).

last_name **parliamentarians last name:** The last name(s) of the parliamentarian, following the cleanup rules specified above.

first_name **parliamentarians first name:** The first name(s) of the parliamentarian, following the cleanup rules specified above.

⁸In case it is known that the MP has no academic title, this variable is set to “none”.

other_name **parliamentarians alternative names or alias:** Contains an (array of) - first or last - name(s) or alias(ses) in case there is or was another first or last name used than mentioned in “first_name” or **last_name**. This might be due to name change, a maiden name, or a commonly used shortage or alias.

gender **parliamentarians gender:** Gives the gender of the parliamentarian in string format. With ‘m’ signifying male and ‘f’ signifying female.

birth_date **parliamentarians birth date:** Mentions the date of birth of a parliamentarian consisting of the two-digit day, the month (first three letters, small, of the English name of the month) and the four-digit year of birth. If only the month is known the the days are dropped and of only the years are know the day and month letters are dropped.

death_date **parliamentarians death date:** Mentions the date of death of a parliamentarian consisting of the two-digit day, the month (first three letters, small, of the English name of the month) and the four-digit year of death. Is 99 if still alive at the moment of data-collection and NA when unknown.

Effort: collect for those who died ‘in office’ and limit to MPs.

citizenship **parliamentarians citizenship(s):** Mentions the parliamentarians - array of - citizenship(s) by mentioning the abbreviation of country(s) he/she is citizen of. The used abbreviations follow the ISO “ALPHA-2” code. (See: http://www.nationsonline.org/oneworld/country_code_list.htm)

Citizenship_Canton **Swiss canton of citizenship:** This variable is Switzerland-specific and contains the abbreviation of the canton(s) in which a parliamentarian holds municipal citizenship(s). If several municipal citizenships exist, the abbreviations are separated by semicolon. Municipal citizenship is inherited from the parents or obtained by way of naturalisation. The same abbreviation is used as in variable “region_abb”

hometown_raw **hometown of a person:** Switzerland-specific string variable. Name (and description) of place/city/municipality that is the candidates/parliamentarians home town.

educ_age Politician its age when finished highest degree (with Diploma).

educ_raw Is a text (string) field containing all available information (in the original fora and language) on the education of a person.

military_string String variable with all information available on the military rank of the politician (in the original form and language).

5 Parliamentary Episode Data Frame (PARE)

The parliamentary episode data-frame contains information that varies across parliamentary episodes. This data-frame connects that parliament (PARL) data-frame with the politician (POLI) data-frame and contains the crucial information of what politicians was in what parliament. Next to that, several time-varying variables are collected for which we know they vary across episodes but on which we do not have more detailed time-stamped information. For example, the number of children. Naturally, this would also be the data-level where a lot of parliamentary behavior variables end up in.

Table 2: Summary of the Parliamentary Episode Level Variables (PARE)

NAME	TYPE	VALUES/EXAMPLE ⁹	SHORT DESCRIPTION	EFFORT
parl_episode_id	PriID	CH_Abate _Fabio.1966 __CH_NT-NR.1946	parliamentary episode ID combination of pers_id and parliament_id	COMP
pers_id	String	CH_Abate _Fabio.1966	See above. A politician is considered a part of a specific parliament as soon as he/she has spend at least one day in the parliament.	COMP

⁹See text below table for more detailed variable specifications

parliament_id	String	[country]-[level] -[abbr].firstyear NL_NT_1946	parliament identifier: See the specification below.	COMP
residence_place_raw	String	“Mecklenburg” “Zuerich” “Essen”	place of residence raw: all available information on the place a parliamentarian lives in	CWA
residence_address	String	“Bismarkstrasse 57, 10627 Berlin”	residence_address: complete residential address of the parliamentarian, only full address otherwise “residence_place_raw” is used	CWE
residence_address_long	String	“Bismarkstrasse 57, 10627 Berlin”	longitude of the residence address	CWE
residence_address_lat	String	“Bismarkstrasse 57, 10627 Berlin”	latitude of the residence address	CWE
family_status_raw	String	‘Widowed’	any (other) family information, in a raw-text format	CWA
married	Binary	1 = ‘married’ 0 = ‘not married’	parliamentarian its marital status; including registered partnerships	CWA
nr_children	Integer	0:15	number of children	CWA

parl_episode_id **unique identifier of the parliament.** Combination of the ...

birth_place_raw **place of birth:** Mentions the place of birth of the parliamentarian. The “place” that is mentioned may be very specific (e.g. street, city) to very broad (e.g. district or state). Which one is mentioned depends on the information available. In case more than one information/level of detail is available, the information is appended to the existing information by a semi-colon. As for the spelling of the names, the same rules apply as for the last_name_full (see “last_name_full”). In case the place has different names in different languages and these are available, the name in the local language is chosen.

residence_place_raw **place of residence:** Mentions the place a parliamentarian lives in. The “place” that is mentioned may be very specific (e.g. city) to very broad (e.g. district or state). Which one is mentioned depends on the information available. In case more than one information/level of detail is available, the information is appended to the existing information by semicolon. In case we have the full address of residence (i.e. street name, house number and city with postcode), the information is recorded in variable “residence_address”. As for the spelling of the place names, the same rules apply as for the last_name_full (see “last_name_full”). In case different place names exist in different languages and are available, the name in the local language is chosen. (see variable “birth_place”)

residence_address **residence_address:** Complete residential address of the parliamentarian. Only full addresses are mentioned (otherwise use the variable “residence_place_raw”). The house number is separated from the street name by a blanc space, the city name and code are separated from the street name and house number by a comma. As for the spelling of the street and city names, the same rules apply as for the last_name_full (see “last_name_full”). In case different names exist in different languages and are available, the name in the local language is chosen (see variable “birth_place”).

married **parliamentarian was married at entry:** Dummy variable capturing whether a parliamentarian was married (or in a registered partnership). Time-varying. If the status is known, but not when then the last parliamentary term is used.

family_status_raw **family status raw:** the raw - in own words - description of one’ relationship status. If the status is known, but not when then the last parliamentary term is used.

nr_children **number of children:** Numeric variable that captures the number of children a parliamentarian has. Time-varying. If the status is known, but not when then the last parliamentary term is used. In the case of conjoined twins, the number of brains are counted.

6 Resume Entries Data Frame (RESE)

What are resume entries? In a nutshell, resume entries refer to the multitude of experiences, activities, and positions that individuals have held throughout their careers. A typical resume entry as used in biographies and CVs contains a description of the activity as well as a start and an end date.

The RESE therefore collects all resume entry variables and variables directly based on or related to these entries.

Table 3: Summary of the Resume Entry Variables (RESE)

NAME	TYPE	VALUES/EXAMPLE ¹⁰	SHORT DESCRIPTION	EFFORT
res_entry_id	String	[country_abb] _[first_last_name] _[first_first_name] _[year_of_birth] _[number] CH_Abate _Fabio.1966__01 DE_Mueller _Johann.1956_dec__06	parliamentarians' resume entry code: a combination of "pers_id" and a count for the number of the entry	COMP
pers_id	String	CH_Abate _Fabio.1966 DE_Mueller _Johann.1956_dec	parliamentarians' individual identification code: a combination of country_abb, the last_name, the first_name and the birth.date	COMP
res_entry_type	Integer	POLI = Political PROF = Professional GEPR = General Prof. EDUC = Educational	parliamentarians' resume entry type: Categorical variable that defines the type of resume entry: Entries can refer to political (1), professional (2), or educational (3) sequences.	COMP
res_entry_start	Date	29apr1976 aug1953 2012	resume entry start date: Variable that captures when a specific resume entry started.	COMP
res_entry_end	Date	29apr1976 aug1953 2012	resume entry end date: Variable that captures when a specific resume entry ended.	COMP
res_entry_at	Date	29apr1976;aug1953 2012	array of at date (array of) dates indicating that this position was held at this point in time.	COMP
res_entry_raw	String	Conseiller communal de 1970 à 1978.	raw entry: The resume entry as is taken from the source.	COMP
res_entry_source	String	Parlement.CH CH Yearbooks	raw entry: The resume entry as is taken from the source. See appendix XX for an overview of used RESE sources.	COMP

¹⁰See text below table for more detailed variable specifications

res_entry_index	Integer	01 02	raw entry: The position of the entry in the source. e.g. 1 means first mentioned. 2 means second etc	COMP
political_function	String	[PrimaryCode1]- [PrimaryCode2]- [SecondaryCode2]- [PrimaryCode3]- [SecondaryCode3]- [PrimaryCode3]- [PrimaryCode4]- [PrimaryCode5]- RE_PA-MA_T1_300_09 NT_LE_T2-CO_1500_03 MU_EX_T3_900_03	political function: Code assigned to the raw entry depending if it is of type “political”.	COMP
parliament_id('s)	String	[country_abb]-[level]-[year] NL_NT_1946 DE_NT_2009 CH_NT_1971 [country_abb]-[level] -[reg_abb]-[year] CH_RE-BS_1997 DE_RE-BW_1956	parliament identifier: Unique code for different parliaments, both across countries/levels and over time. Used if political_function refers to a certain parliament(s). Is an array separated by ' ; ' when there are multiple parliaments	COMP
prof_field	Integer	100:2300	professional field: Professional field codes based on the CAP / U.S. Standard Industrial Classification (cf. level 4 of the political functions primary codes) assigned to the raw entry depending if it is of type “professional”.	
educ_level_isced2011	Integer	1:8	educational level: ISCED level code assigned to the raw entry depending if it is of type “educational”.	COMP
educ_field_isced2011	Integer	00:10	educational field: ISECD field code assigned to the raw entry depending if it is of type “educational”.	COMP

pers_id: Identification code, individual to each parliamentary ACROSS levels. The code consists of the country abbreviation (as specified in variable **country_abb**), the last name (as specified in variable **last_name**). The first name (as specified in **first_name**) and the year of birth (as specified by the last four digits of variable **birth_date**) connected by underscores. In case this id is not unique to one person, the month of birth (as specified in the middle of the **birth_date** variable string) is appended by underscore. If the month of birth is not known or does still not uniquely identify the person, the party abbreviation (as specified by the variable “party_abb”) is appended (additionally to the month of birth) by underscore. Alternative IDs and how they correspond to the main ID can be found in the appendix.

res_entry_id: Identification code for resume entries, individual to each parliamentary and entry. The code consists of “pers_id” and a number per entry. Numbers below 10 are written with an additional zero (e.g. 05) and connected to “pers_id” by two underscores. The order of numbers is arbitrary but corresponds to the order of appearance in the source.

res_entry_type: Categorisation of resume entries. This variable captures the type to which a certain entry can be attributed to. Entries can refer to either political (1), professional (2), or educational (3) biographical

episodes.

res_entry_start: This variable gives the start date of a particular resume entry. These date entries consist of the two-digit day, the month (first three letters, small, of the English name of the month) and the four-digit year of birth. In case of missing information on days or month, shortened versions of the date are used.

res_entry_end: This variable gives the end date of a particular resume entry. These date entries consist of the two-digit day, the month (first three letters, small, of the English name of the month) and the four-digit year of birth. In case of missing information on days or month, shortened versions of the date are used.

res_entry_raw: This variable gives the raw resume entry string as it has been extracted from the original source.

political_function: Multi-level code that captures various facettes of resume entries that are categorised as political under “res_entry_type”. A thorough presentation of the political functions coding scheme is provided in the subsequent subchapter.

parliament_id: Unique identification code for different parliaments, both across countries/level and over time. It is added whenever ‘‘political_function’’ refers to a function connected to national and regional parliaments. Is an array separated by ‘;’ when there are multiple parliaments.

prof_field: This variable captures the professional area of resume entries that were categorised as professional under “res_entry_type”. Professional field codes are based on the CAP / U.S. Standard Industrial Classification. The coding scheme corresponds to the policy areas (level 4) of the political functions primary codes. An overview of the codes is provided in the subsequent subchapter.

educ_level_isced2011: Ordinal variable capturing the level of education of particular resume entries. Based on the ISCED 2011 international coding standard, only first digit is coded. see <http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf> for the codes.

isced2011 levels

- ISCED level 0 – Early childhood education
- ISCED level 1 – Primary education
- ISCED level 2 – Lower secondary education
- ISCED level 3 – Upper secondary education
- ISCED level 4 – Post-secondary non-tertiary education
- ISCED level 5 – Short-cycle tertiary education
- ISCED level 6 – Bachelor’s or equivalent level
- ISCED level 7 – Master’s or equivalent level
- ISCED level 8 – Doctoral or equivalent level

educ_field_isced2013: The isced field classification see <http://www.uis.unesco.org/Education/Documents/isced-fields-of-education-training-2013.pdf> only the first digit is coded.

isced2013 fields

- ISCED Field 00 - Generic programmes and qualifications
- ISCED Field 01 - Education
- ISCED Field 02 - Arts and humanities
- ISCED Field 03 - Social sciences, journalism and information
- ISCED Field 04 - Business, administration and law
- ISCED Field 05 - Natural sciences, mathematics and statistics
- ISCED Field 06 - Information and Communication Technologies
- ISCED Field 07 - Engineering, manufacturing and construction
- ISCED Field 08 - Agriculture, forestry, fisheries and veterinary
- ISCED Field 09 - Health and welfare
- ISCED Field 10 - Services

6.1 Political functions

6.1.1 Principles for Classification of Political Functions

Working definition: Political functions are functions that are directly or indirectly aimed at creating and shaping policy. The proposed coding scheme for the classification of political functions consists of two parts. The first one is labelled **primary political function coding** while the second one can be referred to as **secondary political function coding**.

- **Primary political function coding** entails information that can be coded for all political functions. For example, we should be able to code all political functions with regard to the level they are located on in the multi-level hierarchy.
- **Secondary political function coding** refers to information that is specific to subsets of the primary coding. With the primary coding scheme for instance, we are unable to distinguish whether some-one's party function was in a youth party (1) or not (0). Secondary codings will therefore be introduced whenever an even more detailed distinction is of interest for particular research questions.

6.1.2 Primary Political Function Coding

Level 1: Position of organisation in the multi-level hierarchy Coding the position in the multi-level structure depends on the institutional affiliation of the political function (thus often answering the question *who pays you?*). Example: Working as an ambassador to the EU will therefore be coded as national level (C).

Effort: collect until complete.

- IN - International level
- EU - EU level
- NT - National level
- SR - Superregional level (regional, super) (below national, above regional level)
- RE - Regional level (NL: provinces; DE: Bundesländer; CH: Cantons)
- SD - Superdistrictual level (districutal, super) (below regional, above district level)¹
- DI - District level (CH: Ämter, Bezirke; DE: Kreise)
- SM - Supermunicipal level (municipal, super) (below district, above municipal level; CH: Kreise)¹
- MU - Municipal level

Level 2: Organisational type -

Effort: collect until complete.

- LE - Legislative
- EX - Executive
- JU - Judiciary
- AD - Administration (including dependent agencies)
- SI - Semi-independent agencies
- PA - Party
- IG - Interest Group
- OT - Other

Note: if you are a CEO of a large company and it is part of your job to influence policy you would be coded as organisational type: IG - Interest Group

Semi-independent agencies are public bodies that resemble administrations in their behaviour. They have goals defined by the state such as financial regulation (e.g. FINMA), drug licencing (SWISSMEDIC), or nuclear power safety (ENSI). Furthermore, they can also include state-owned research institutions (e.g. AGROSCOPE). They do not pursue financial goals. Publically owned companies (e.g. Swiss Federal Railways, Swisscom, Swiss Post) are therefore NOT semi-independent agencies but IGs.

Level 3: Internal structure of organisational types, up to 3 sub-levels -

Effort: collect when available.

- T1 Top-tier
- T2 Mid-tier
- T3 Low-tier

How to code the tiers The terms *top-tier*, *mid-tier*, *low-tier* would be used as overarching descriptions. For the specific combinations of level 1 and 2 codes (e.g. B-A EU-level legislative), the name of the actual corresponding tier name would be used e.g. *parliamentary presidency* instead of *top-tier*.

Institutions have highly differing structures that do not only vary across but even within countries. The use of three tiers is necessary to enable comparisons between institutions within and across countries. The tiers offer a meaningful simplification of institutional structure in terms of policy competences. Almost all institutions depict a similar structure: the highest and leading instance holding the broadest/deepest competences (T1), the unit right below the highest instance which holds specialised competences and reports to or is controlled by the highest instance (T2), and the rest of the organisational hierarchy which covers the remaining less powerful and less prestigious functions (T3).

Because parties and interest groups have non-standardised organisational structure, we will only

It should be noted that the tiers do not (necessarily) refer to a division of labour though. Support bodies tasked with administrative and implementing functions (e.g. secretariats affiliated with leadership or steering bodies) are coded on the same tier as their principals'. However, bodies to which policy competences are delegated are coded on a separate level.

As can be seen in the table below almost all institutions under consideration either follow such a three tier structure or the structure can be meaningfully divided into three tiers.

Code	Description	Categories	Examples
IN-...	International level
EU_LE	EU-level legislative	<ol style="list-style-type: none"> 1. Parliamentary presidency 2. Committee / delegation / party group 3. Assembly 	<ol style="list-style-type: none"> 1. "Schulz' cabinet" 2. "Committee on Legal Affairs" 3. "MEP"
EU_EX	EU-level executive	<ol style="list-style-type: none"> 1. Commission presidency 2. College of Commissioners 3. Secretariat-General 	<ol style="list-style-type: none"> 1. "First Vice-President" 2. "Commissioner for Energy Union" 3. "Unit B3 Ethics"
EU_JU	EU-level judiciary	DNC <ol style="list-style-type: none"> 1. European Court of Justice 2. General Court 3. Specialised court 	<ol style="list-style-type: none"> 1. "Third Chamber" 2. "Registry" 3. "EU Civil Service Tribunal"
EU_AD	EU-level administration	<ol style="list-style-type: none"> 1. Directorate-General 2. Directorate / department 3. Unit 	<ol style="list-style-type: none"> 1. "DG Communication" 2. "Directorate Resources" 3. "Energy policy coordination unit"
EU_SI	EU-level semi-independent agencies	<ol style="list-style-type: none"> 1. Decision-making body 2. Department 3. Unit 	<ol style="list-style-type: none"> 1. "ECB's executive board" 2. "Department of air and climate change" 3. "Nutrition unit"
EU_PA	EU-level party	<ol style="list-style-type: none"> 1. Party leadership body 2. Steering body (commission, delegation, council, 'bureau', non-administrative secretariat) 	<ol style="list-style-type: none"> 1. "Presidency" 2. "PES Council"
EU_IG	EU-level interest group	<ol style="list-style-type: none"> 1. Leadership body 2. Steering body / department (committee, delegation, council, bureau, secretariat, office) 	<ol style="list-style-type: none"> 1. "Board", "presidential council", "Directorate-General" 2. "General committee", "department", "committee", "working groups", "Brussels office"
EU_OT	Other EU level entities

NT_LE	National parliament	<ol style="list-style-type: none"> 1. Parliamentary presidency 2. Committee / delegation / party group 3. Assembly 	<ol style="list-style-type: none"> 1. “Nationalratspräsidentin” 2. “Commissie voor de Rijksuitgaven” 3. “MP”
NT_EX	National government	<ol style="list-style-type: none"> 1. Government Leadership 2. Ministerial Level 3. Junior Ministerial Level 	<ol style="list-style-type: none"> 1. “Bundeskanzlerin” 2. “Bundesminister des Innern” 3. “Staatssecretaris van Onderwijs”
NT_JU	National judiciary	DNC <ol style="list-style-type: none"> 1. Highest court 2. Court of Appeal / 2nd highest court 3. - 	<ol style="list-style-type: none"> 1. “Raad van State”; Hoge Raad“, “Bundesfinanzhof“, “Bundesgericht” 2. “Bundesverwaltungsgericht”
NT_AD	National administration	<ol style="list-style-type: none"> 1. Department 2. Subdepartment 3. Unit 	<ol style="list-style-type: none"> 1. “Secretariaat-generaal“, “Generalsekretariat“, “Staatssekretariat” 2. “Bundesamt (CH)“, Abteilung (DE)“, Directie-Generaal“ 3. Referat“, Stab“, Unterabteilung“, Amt“ Directie“
NT_SI	National semi-independent agencies	<ol style="list-style-type: none"> 1. Decision-making body 2. Department 3. Unit 	<ol style="list-style-type: none"> 1. “Bestuur“, “Direktion (CH)” 2. “Directie“, “Abteilung“ 3. “Team (NL)“, Referat (DE)“
NT_PA	National party	<ol style="list-style-type: none"> 1. Party leadership body 2. Steering body (commission, delegation, council, bureau, secretariat) 	<ol style="list-style-type: none"> 1. “Partijbestuur“, Parteileitung“ 2. Partijbureau“, “Generalsekretariat“, “Bundesgeschäftsstelle“, Bundesparteigericht“
NT_IG	National interest group	<ol style="list-style-type: none"> 1. Leadership body 2. Steering body / department (committee, delegation, council, bureau, secretariat, 	<ol style="list-style-type: none"> 1. “Board“, “presidential council“, “Directorate-General” 2. “General committee“, “department“, “committee“, “working groups“, “Brussels office”

NT_OT	Other national entities	...	e.g. intercantonal concordats
SR_	Superregional level	...	CH: 11 Kreispostdirektionen (bis 1997), 3 Kreise der Bundesbahnen (bis 1998), 4 Kreise der Zollverwaltung (bis 2000), 14 Kreise der Alkoholverwaltung NL: 5 gerechtshoven (courts of appeal in five regions)
RE_LE	Regional parliament	<ol style="list-style-type: none"> 1. Presidency of the regional parliament 2. Committee / delegation / faction 3. Assembly (of the regional parliament) 	<ol style="list-style-type: none"> 1. "Parlamentspraesident", "Kantonsratspraesident", "Grossratspraesidentin", Landtagspraesident" 2. "Ausschuss fuer Laendlichen Raum und Verbraucherschutz", "Presidium", "Fractievoorzittersoverleg" 3. "Provinciale Staten", Landtag", Kantonsrat", Grossrat"
RE_EX	Regional government	<ol style="list-style-type: none"> 1. Regional Government Leadership 2. Regional Ministerial Level 3. Regional Junior Ministerial Level 	<ol style="list-style-type: none"> 1. "Ministerpraesident", "Comissaris van de Koning", "Regierungspraesident", Staatsratspraesidentin", 2. "Gedeputeerde Staten", "Regierungsrat", "Staatsrat", Innenminister" 3. ...
RE_JU	Regional judiciary	<ol style="list-style-type: none"> 1. Highest Court 2. Court of Appeal/ 2nd highest court 3. - 	<ol style="list-style-type: none"> 1. "Oberlandesgericht Karlsruhe" (highest ordinary court) "Rechtbank", "Kantonsgericht" 2. "Amtsgericht Achern"
RE_AD	Regional administration	<ol style="list-style-type: none"> 1. Department 2. Subdepartment 3. Unit 	<ol style="list-style-type: none"> 1. "Direktion" 2. "Kantonsamt (CH)", Abteilung (DE)", Directie-Generaal" 3. Referat", Stab", Unterabteilung", Amt" Directie"
RE_SI	Regional semi-independent agencies	<ol style="list-style-type: none"> 1. Decision-making body 2. Department 3. Unit 	<ol style="list-style-type: none"> 1. "Bestuur", "Direktion (CH)" 2. "Directie", "Abteilung" 3. "Team (NL)", Referat (DE)"

RE-PA	Regional party	<ol style="list-style-type: none"> 1. Party leadership body 2. Steering body (commission, delegation, council, bureau, secretariat) 	<ol style="list-style-type: none"> 1. “Partijbestuur“, Landesvorstand“, Landespraesidium“, Kantonalparteileitung“ 2. Partijbureau“, Landesgeschäftsstelle“, Landesparteigericht“, “Parteisekretariat”
RE_IG	Regional interest group	<ol style="list-style-type: none"> 1. Leadership body 2. Steering body / department (committee, delegation, council, bureau, secretariat, 	<ol style="list-style-type: none"> 1. “Board”, “presidential council”, “Directorate-General” 2. “General committee”, “department”, “committee”, “working groups”, “Bremen office”
RE_OT	Other regional entities	...	e.g. private-public partnerships with political functions
SD_	Superdistrictual level	...	NL: 19 rechtbanken (separated into judicial arrondissements), waterschappen CH: Zivilkreisgerichte
DI_LE	District parliament	<ol style="list-style-type: none"> 1. Presidency of the district parliament 2. Committee / delegation / faction 3. Assembly (of the district parliament) 	<ol style="list-style-type: none"> 1. “Kreistagspraesident“, “Landrat” 2. “Ausschuss fuer Migration” 3. “Kreistagsabgeordnete”
DI-EX	District government	<ol style="list-style-type: none"> 1. District government leadership 2. District government membership 3. District government steering body 	<ol style="list-style-type: none"> 1. “Landrat” “Landrat“, “Amtsstatthalter“, “Regierungsstatthalter“, Bürgerworthalter“ 2. “Kreisvorstand” 3. ...
DI-JU	District judiciary	<p>dnc</p> <ol style="list-style-type: none"> 1. District Level Court 	<ol style="list-style-type: none"> 1. “Amtsgericht Achern“, “Kreisgericht“, “Amtsgericht“, “Bezirksgericht“, Kantongerecht“
DI-AD	District administration	<ol style="list-style-type: none"> 1. Department 2. Subdepartment 3. Unit 	<ol style="list-style-type: none"> 1. “Kreisamt (CH)“, “Dezernat (DE)“, “Abteilung (DE)“, 2. “Referat“, “Stab” 3. “Unterabteilung”

DI-SI	District semi-independent agencies	1. Decision-making body 2. Department 3. Unit	...
DI-PA	District level party	1. Party leadership body 2. Steering body (commission, delegation, council, bureau, secretariat)	1. "Partijbestuur", Präsidium", Kamerkringbestuur", "Geschäftsleitung", Bezirksvorstand", Kreispraesidium", Partijbureau", Kreisgeschäftsstelle", 2. Kamerkring", Kamercentrale", Provinciaal contact"
DI-IG	District level interest group	1. Leadership body 2. Steering body / department (committee, delegation, council, bureau, secretariat,	1. "Board", "presidential council", "Directorate-General" 2. "General committee", "department", "committee", "working groups", "Bremen office"
DI-OT	Other district level entities e.g. private-public partnerships with political functions
SM-	Supermunicipal level	...	CH: "Gemeindeverband"
MU-LE	Municipal parliament	1. Presidency of the municipal parliament 2. Committee / delegation / faction 3. Assembly	1. "Parlamentspraesident" 2. "Ausschuss" 3. Gemeindeparlament", "Einwohnerrat", "Stadtparlament", "Gemeenteraad"
MU-EX	Municipal government	1. Municipal government leadership 2. Municipal government membership 3. Municipal government steering body	1. Gemeindepräsident", Stadtverordnetenvorsteher", Ratsvorsitzender", Präsident der Bürgerschaft", Bürgervorsteher", (Ober-)Bürgermeister", Burgemeester" 2. Gemeinderat", Stadtrat", College van burgemeester en wethouders" 3. "Welstand", "SER", "Einbürgerungskommission"
MU-JU	Municipal judiciary	DNC 1. (see District Judiciary)	...

MU-AD	Municipal administration	<ol style="list-style-type: none"> 1. Department 2. Subdepartment 3. Unit 	<ol style="list-style-type: none"> 1. “Ressorts”, “Amt”, “Dezer-nat” 2. “Bereich” 3. “Sachbereich”
MU-SI	Municipal semi-independent agencies	<ol style="list-style-type: none"> 1. Decision-making body 2. Department 3. Unit 	...
MU-PA	Municipal party	<ol style="list-style-type: none"> 1. Party leadership body 2. Steering body (commission, delegation, council, bureau, secretariat) 	<ol style="list-style-type: none"> 1. “Partijbestuur“, Vorstand des Ortsverbandes“, “Geschäftsleitung“, Präsidium“, “Geschäftsleitung“ 2. Partijbureau“, Parteibüro“
MU-IG	Municipal interest group	<ol style="list-style-type: none"> 1. Leadership body 2. Steering body / department (committee, delegation, council, bureau, secretariat, 	<ol style="list-style-type: none"> 1. “Board” “presidential council”, “Directorate-General” 2. “General committee”, “department”, “committee”, “working groups”
MU-OT	Other municipal entities

Table 4: Political functions level 3 coding specifications

While this three-tier code condenses information, the more detailed information will not necessarily be lost. On the one hand, it is still part of the string variable that is used to code the tier; on the other hand specific secondary codes can be introduced. The secondary code (see 2 Secondary Coding) can be used to introduce more specificity whenever it is required. Take for example national level parliaments and say we are particularly interested in positions in national parliamentary bodies. The primary coding does not provide detailed information on that:

Code	Description	Categories
NT-LE	National parliament	<ol style="list-style-type: none"> 1 - Parliamentary presidency 2 - Committee / party group / delegation 3 - Assembly

Table 5: Example of Three-tier Coding

The information on committees, delegations, and party groups is condensed into one single category (mid-tier). If the interest arises to study membership in these different bodies separately, the secondary code would capture that (See table 6)

Code	Description	Categories	Secondary code options for NT-LE-02
NT-LE	National parliament	1 - Parliamentary presidency 2 - Committee / delegation / party group 3 - Assembly	1: Committee 2: Faction (party group) 3: Parliamentary delegation

Table 6: Example of Secondary Coding

Beyond the coding of just NT-LE-02 for national parliamentary bodies, we can thus in the same step code this additional information.

Level 4: 21 policy areas (taken from the Comparative Agendas Project and extended by areas of economic activities taken from the U.S. Standard Industrial Classification). Only used if applicable/available. These categories are not mutually exclusive (e.g. functions can be coded into multiple policy areas). Information for policy area may not always be available or applicable (e.g. municipal councillors dealing with a wide range of policy issues). Whenever none of the subcategories (sub-areas) apply, the main topic (policy area) is to be coded.

-

Effort: collect when available.

Code	Policy Area	Sub-Area	Notes
(0)100	Domestic Macroeconomic Issues		
(0)200	Civil Rights, Minority Issues, and Civil Liberties		
(0)201		Ethnic Minority and Racial Group Discrimination	
(0)202		Gender Discrimination	Given particular research interests in gender issues, the aspect of sexual orientation discrimination has been dropped from the original CAP code.
(0)300	Health		All health related activities and entities are coded here. This includes aspects related to health facilities, insurance, pharmaceutical products, tobacco, health research, and the health care system in general.
(0)400	Agriculture		
(S0)401		Crop and livestock	Based on the (U.S.) Standard Industrial Classification SIC 0100 and 0200. This category covers aspects related to the production poultry , livestock, dairy, fruit, nuts, and crop plants.
(S0)402		Forestry	Based on the (U.S.) Standard Industrial Classification SIC 0800. If forestry is related to Forest Protection, then it is coded as 700 Environment.
(0)403		Agricultural Primary Processing	This category is not based on any CAP or SIC code. It covers aspects related to primary processing of poultry , livestock, dairy, fruit, nuts, and crop plants. These are the steps of food processing directly after production. Wine and wineries are also coded here. This category is also given precedence over S1505 Manufacturing of Food and Kindred Products whenever a clear connection to agriculture is predominant in later food processing steps.

(0)404		Agricultural Trade, Retail, Services, and Supply	This category is not based on any CAP or SIC code. It covers all activities related to wholesale and retail trade of agricultural products, services and supply provided for / or by agriculturally focussed entities.
(0)405		General Agricultural Activities	This category is not based on any CAP or SIC code. It covers all activities related to agriculture that are not in anyway covered by the other agriculture categories. This includes for instance agricultural interest groups, farmers' newspapers and mountain farmers' relief organisations, agricultural disaster insurance, agricultural research institutions, animal welfare, pest control etc.
(0)406		Fisheries, Fishing, and Hunting	Based on CAP code 408.
(0)500	Labour and Employment		E.g. Functions in pension funds
(0)600	Education		
(0)700	Environment		
(0)800	Energy		
(0)900	Immigration and Refugee Issues		
1000	Transportation		Everything environment-related is coded under 700 Environment. Transportation that serves primarily a touristic purpose should be coded under S1515.
1200	Law, Crime, and Family Issues		
1300	Social Welfare		
1400	Community Development and Housing Issues		
(S)1401		Real Estate	Partially based on the (U.S.) Standard Industrial Classification SIC 7000.
1500	Economic Activities and Domestic Commerce		This category encompasses a wide array of economic activities which do not fall under the scope of the other 20 policy areas. When economic activities come with interactions with public officials, they are coded according to their economic area. Note that activities covered by other policy areas such as farming (400 Agriculture) or the production of pharmaceutical goods (300 Health) are coded in the corresponding policy areas.
(S)1501		Mining and Quarrying of Metals and Nonmetallic Minerals, except Fuels	Based on the (U.S.) Standard Industrial Classification SIC 1000 and 1400. Energy resources (oil, gas, coal etc.) are coded under 800 Energy.
(S)1502		Construction, Engineering, Architecture, and Surveying	Based on the (U.S.) Standard Industrial Classification SIC 1500-1799 and 8710

(S)1503		Secondary Processing of Food and Kindred Products	Partially based on the (U.S.) Standard Industrial Classification SIC 2000. This category comprises secondary processing of food products. Entities active in this area (often) do not execute primary processing OR produce mainly for end customers. Hence, agricultural policies are mostly NOT the main concern of these actors active in this area. Non-grape based alcoholic beverages are coded here.
(S)1504		Manufacturing of Leather, Textile Products, and Apparel	Based on the (U.S.) Standard Industrial Classification SIC 2200, 2300, 3100.
(S)1505		Manufacturing of Lumber and Wood Products, incl. Furniture and Paper	Based on the (U.S.) Standard Industrial Classification SIC 2400, 2600, 3500. Activities related to unprocessed timber are coded under S402 Forestry.
(S)1506		Manufacturing of Chemicals and Allied Products	Based on the (U.S.) Standard Industrial Classification SIC 2800. Pharmaceutical companies are coded under 300 Health.
(S)1507		Manufacturing of Rubber and Miscellaneous Plastics Products	Based on the (U.S.) Standard Industrial Classification SIC 3000.
(S)1508		Manufacturing of Stone, Clay, Glass, and Concrete Products	Based on the (U.S.) Standard Industrial Classification SIC 3200.
(S)1509		Manufacturing of Primary Metal Products	Based on the (U.S.) Standard Industrial Classification SIC 3300.
(S)1510		Manufacturing of Machinery, Computer Equipment, Electronic and other Electrical Equipment, Measuring, Analysing, and Controlling Instruments, Photographic, Medical and Optical Goods; Watches and Clocks.	Based on the (U.S.) Standard Industrial Classification SIC 3400-3600, 3800. Equipment related to transportation (motor vehicles, aircrafts, railroad equipment, ships, bicycles, motorcycles etc.) are coded under 1000 Transportation. Manufacturing of equipment mainly for agricultural purposes / agricultural customers is coded under 404 Agricultural Trade, Retail Services, and Supply.
(S)1511		Wholesale Trade	Based on the (U.S.) Standard Industrial Classification SIC 5000-5199. Wholesale directly tied to agricultural organisations is coded under 404 Agricultural Trade, Retail Services, and Supply.

(S)1512		Retail Trade	Based on the (U.S.) Standard Industrial Classification SIC 5200-5999. This category includes activities related to General Merchandise Stores, Food Stores, Building Materials, Hardware, Garden Supply, and Mobile Home Dealers, Apparel and Accessory Stores, and Home Furniture, Furnishings, and Equipment Stores. Automotive Dealers and Gasoline Service Stations are coded under 1000 transportation. Retail directly tied to agricultural organisations is coded under 404 Agricultural Trade, Retail Services, and Supply.
(S)1513		Banks, Financial Institutions, and Non-Health / Non-Accident Insurances	Based on the (U.S.) Standard Industrial Classification SIC 6000-6200 and 6400. Also includes 8720 (Accounting, Auditing, and Bookkeeping Services) and 8740 (Management and Public Relations Services). Health and accident related insurances (6320) are coded under 300 Health.
(S)1514		Tourism, Hotels, and Restaurants	Based on the (U.S.) Standard Industrial Classification SIC 7000 and 5800
(S)1515		Other Services	Based on the (U.S.) Standard Industrial Classification SIC 7000-8999. This category includes a wide array of services including Personal and Business Services, Amusement and Recreation Services (if S1515 [Tourism, Hotels, and Restaurants] and 2300 [Cultural Policy Issues] are not applicable), Legal Services, and Social Services (e.g. child day care).
1516		Interbranch and Cross-Sectoral Activities	Economic activities that span multiple industries, sectors, and/or services where no clear emphasis on a particular area of business activity emerges. Examples include holding companies and general business associations.
1600	Defense		
1700	Space, Science, Technology, and Communications		
(S)1701		Research, Development, and Testing	Based on the (U.S.) Standard Industrial Classification SIC 8730.
1702		Newspaper, Publishing, and Broadcast Industry Regulation	Dissemination of information related to any other policy area (e.g. publishing a farmers' newspaper) is coded in the corresponding policy area. Based on CAP code 1707.
1703		Computer Industry and Computer Security	Based on CAP code 1709.
1800	Foreign Trade		
1900	International Affairs and Foreign Aid		
2000	Government Operations		

2100	Public Lands, Water Management, and Territorial Issues		
2300	Cultural Policy Issues		Also includes sports, heritage, and history
2400	General Ideological Issues		Only applicable to interest groups. Captures general ideological stances (e.g. liberalism, conservatism, socialism) of organisations that do not compete for public office

Table 7: Political functions level 4 coding specifications: policy area codes Note: The letter S in front of the numerical code specifies newly created codes based on the U.S: SIC classification.

Level 5: Occupational position on the hierarchical level at hand (categorical scale!) *Effort: collect when available.*

The codes capture occupational position in a standardised form. Every piece of information will be assigned a numerical value (e.g. the president of an organisational unit will be coded as leadership [01]). More detailed coding can again be introduced by secondary coding (e.g. leadership in parties could be split into president (1) and vice-president (2); see the example above for committee / delegation / faction)

Code	Category	Examples
01	Leadership	“Präsident”, “Vorsitzende”, “Teamleiterin”, “Chair”, “Co-président”, “Co-Direktor”, “Gesellschafter”
02	Deputy leadership	“Vice-chair”, “stv. Amtsleiter”
03	Ordinary worker / member	“Member of the council”, “Mitarbeiter”, “Erstatzmitglied”
04	Leading support staff ¹¹ to leadership	“Head of the party bureau”
05	Support staff to leadership	“Personal assistant to the secretary-general”
06	Leading support staff	“Head assistant for MEP”
07	Support staff	“Assistant for MEP”, “secretary”
09	Active (position not specified)	“Amt für Volksschulbildung 1974-9”

Table 8: Level 5: Occupational Position Codes

Distinguishing Level 3 and 5: A Short Note on the Difference Between Leaders and Members

The distinction between the internal structure of organisations and the occupational position can be a cause for confusion. Depending on whether a position is considered over the entirety of the organisation or merely the tier (1, 2, or 3) at hand, the level 5 code may vary. However, only the second option is correct. Occupational positions are always only coded in relation to their tier.

To illuminate this issue more, let us consider an example: how do we code ministers? It is quite clear that ministers are located on tier 2 (i.e. on the ministerial level; _T2_). They are neither government leadership (tier 3; _T1_) nor located on the junior ministerial level (_T3_). The more difficult question is then, however, how to code their occupational position (level 5). In relation to the head of government (_T1_NA.01), one might **incorrectly** assume that ministers are ordinary workers / members. They are not though. Rather, they are considered the leadership of their own tier (i.e. tier 2). Hence ministers are always coded as (_T2.01).

Figure 2 provides an overview of several political functions in the executive and what code they are assigned. The key element to this overview is that the distinction between leadership and other functions is always made at the tier at hand, not across tiers.

In analogy, the same applies to the legislative. Positions in parliament are always coded only with reference to the tier at hand. This might seem obvious for some positions and counterintuitive for others: The parliamentary president is naturally coded as the leader (01) of the highest level in parliament (T1) i.e. as _T1_NA.01. Similarly, members of committees, parliamentary party groups, and delegations are coded as members of their respective bodies (_T2_NA.03).

On every tier, the highest position is **always** coded as leadership. Confusion might arise when the distinction between leadership and ordinary members is concerned on the assembly level (T3). Members of parliament are always as leaders because they are the highest instance on their level. Leadership, in that sense, thus does not only refer to exercising power but also relates to independence in making one’s decisions.

Figure 3 provides an overview of several political functions in the legislative and what code they are assigned.

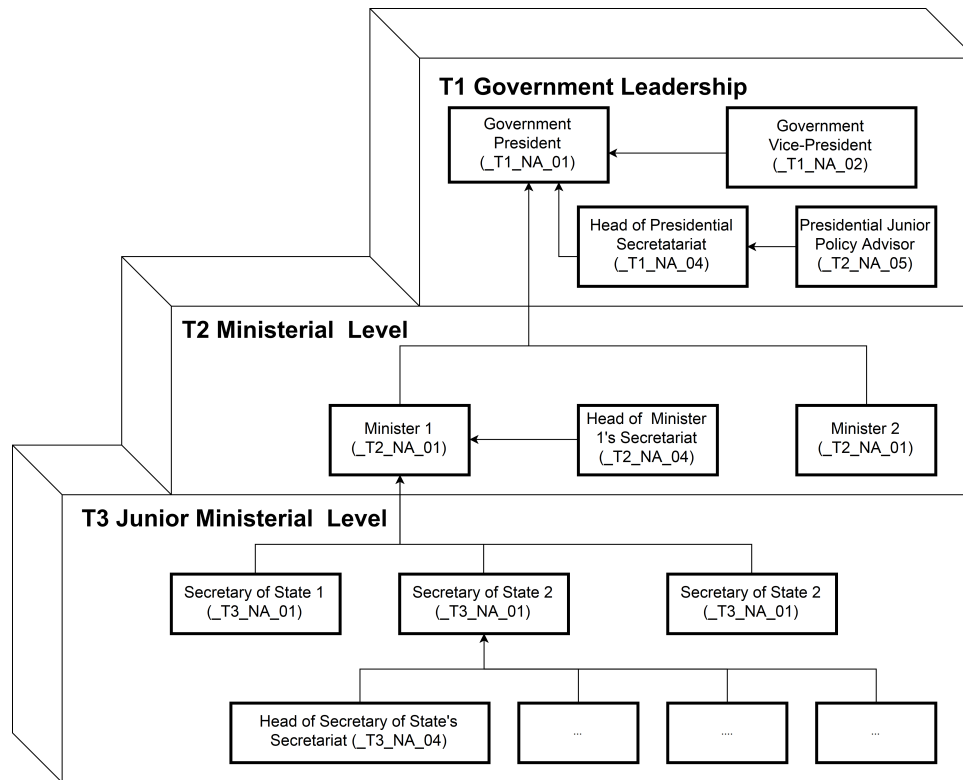


Figure 2: Overview of political functions in the executive and how they are coded on level 3 and 5.

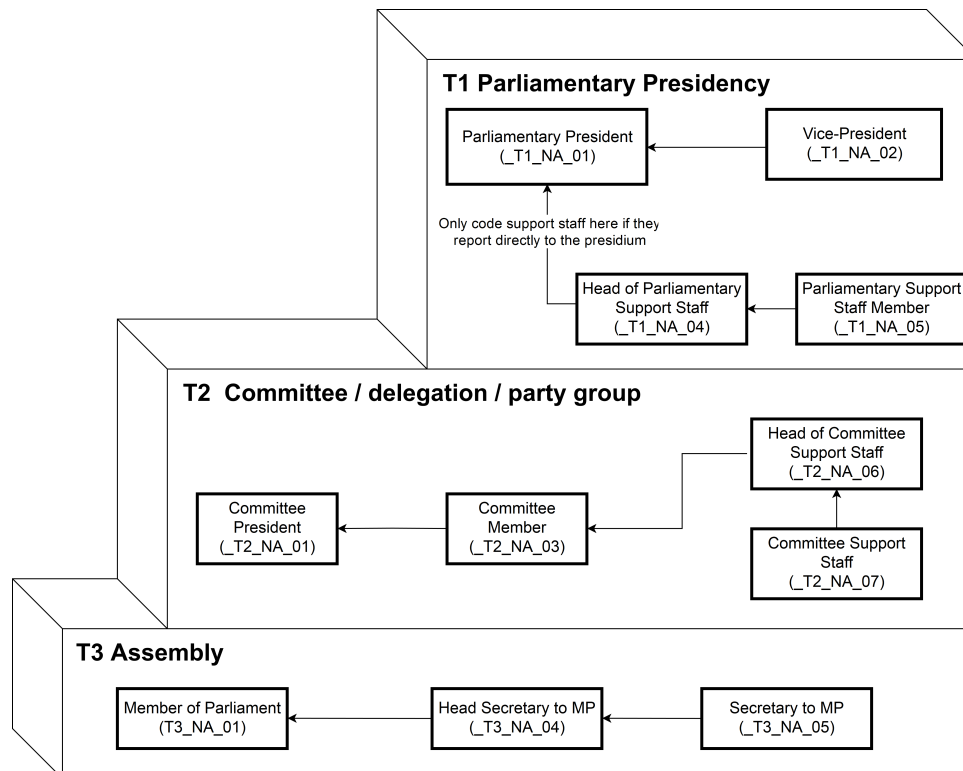


Figure 3: Overview of political functions in the legislative and how they are coded on level 3 and 5.

6.1.3 Secondary Political Function Codes

Secondary coding's are research project (e.g. research question) specific 'even more detailed' distinction in the coding of political functions. We are happy to facilitate you with the coding of for additional secondary codes.

Drop us a message at info@parliamentarycareersincomparison.org The secondary codes in the next paragraphs are already available.

Secondary Political Function Coding: Elected or Non-Elected Applies to: ‘X’_LE_ ‘X’_EX_: Whenever the organisational type of a political function is specified as either legislative (level 2 = LE) or executive (level 2 = EX) or judicial (JU), then it will be specified if this functions was **elected** or **non-elected**.

Code	Category	Examples
PE	Probably ¹² elected	“Minister”, “MP”
NE	Non-Elected	“Assistant to the Minister”, “Parliamentary Committee Secretary”

Table 9: Elected or Non-Elected Political Functions - **DNC**

Secondary Political Function Coding: Legislative steering body specification Detail on the second tier of all legislative (X_LE) positions is important. Hence a further distinction between the three types of steering bodies is made as specified below.

Applies to: ‘X’_LE_T2-: Whenever the organisational type of a political function is specified as a legislative body and the tier as T2 (level 2 = LE, and level 3 = T2), then more detailed information on the type of legislative body is collected:

Code	Category	Explanation / notes	Example
CO	Committee	Legislative committee	“Committee on Legal Affairs”, “Kommission für Verkehr und Fernmeldewesen des Nationalrates”
SC	Subcommittee / working group	Subcommittee and working groups of legislative committee	“Subkommission EDA/VBS”, “Unterausschuss Auswärtige Kultur- und Bildungspolitik”, “Arbeitsgruppe Hochseeschiff-fahrtsbürgschaften”
DL	Delegation	Legislative delegation	“Delegation bei der parlamentarischen Versammlung der NATO”, “Deutsche Delegation in der OSZE”
FA	Faction party group	Legislative faction	“Grüne Fraktion”, “Groenlinks”
AG	All-Party Parliamentary Groups	All-Party Parliamentary Groups (APPGs) are informal topical cross-party groups that have no official status within Parliament.	“Parlamentarische Gruppe für Altersfragen”, “Parlamentarische Gruppe Binnenschifffahrt”

Table 10: Specification of political function secondary code: legislative steering body codes

Secondary Political Function Coding: Advisory ‘extra parliamentary’ expert committee Applies to: ‘X’_AD_T2-: Whenever the organisational type of a political function is specified as a administrative body (level 2 = AD), then we have to option to specify the membership of special expert committees. E.g. ‘committee installed by the Minister’ in the Netherlands.

Secondary Political Function Coding: Youth versus Regular party Applies to: ‘X’_PA_: Whenever the organisational type of a political function is specified as a party (level 2 = PA), then more detailed information on the type of party is collected:

Code	Category	Examples
MA	Main party	“FDP”
YO	Youth party	“JUSO”

Table 11: Youth or Regular Party Political Functions

Secondary Political Function Coding: Regular versus Shadow positions - DNC (Table 12) Applies to: ‘X’_LE/EX]-: Whenever the organisational type of a political function is either legislative or executive, then it will be specified if this function is the regular or the shadow function.

Code	Category	Examples
01	Regular position	
02	Shadow position	“Official Opposition Shadow Cabinet”

Table 12: Shadow or Regular Political Functions - DNC

Secondary Political Function Coding: Important party committees (Table 13) Applies to: ‘X’_PA_T2: Whenever the organisational type of a political function is specified as a party (level 2 = PA) steering body / committee (level 3 = 02), then more detailed in-formation on the type of steering body is collected. Please note that steering bodies can exist out of one or more individuals.

Code	Category	Examples
10	Substantive committee per policy area. See level 4! (6.1.2)	
20	Election specific committees	
_21	(_Election) programme committee	“Wahlprogrammkommission”
_22	_Election campaign team	“Wahlkampfkommission”
_23	_Candidate selection committee	
_24	_Referendum / popular initiative campaign team (CH)	
_25	_Other Election Committee	
30	Other organisational (not policy-area specific) committees	
_31	_Ideological faction	“Seeheimer Kreis”
_32	_Regional committee	“Forum Ostdeutschland der Sozialdemokratie”
_33	_Training/development/recruitment	
_34	_Party magazine	
_35	_Other organisational committee	
40	External party-controlled organisations	Note: Need to be legally independent
_41	_General scientific bureau	
_42	_Fundraising body	

Table 13: Important Party Committees

Secondary Political Function Coding: Interest groups codes Closely based on the typology of the <http://interarena.dk/> project. Applies to: ‘X’_IG-: Whenever the organisational type of a political function is specified as an interest group (level 2 = IG), then more detailed information on the type of interest group has to be collected.

Code	Category	Explanation / notes	Example
10	Unions	Associations of employees negotiating work-related terms and conditions	“Blue-collar union”, “Federatie Nederlandse Vakbeweging”
20	Business org.		
21	Business / trade association	Associations of firms	“Verband der Chemischen Industrie e.V. (VCI)“, “Arbeitgeberverband der Schweizer Maschinenindustrie“
22	Private Firm / company	Firms by majority in possession of private owners. Firms are only code-able as interest groups if the position within the firm falls mainly into the area of public affairs.	“Board of oversight at Shell Global“
23	Public Firm / company	Firms by majority in possession of public owners (state, regions, municipalities etc.).	“Board of oversight at SBB AG“

30	Institutional or- ganisation	Organisation of public authorities or institutions that are primarily publically funded	“swiss universities”, Netz der Berner Spitäler und Kliniken”
40	Occupational or- ganisation	Organisation of employees not negotiating terms and conditions	“Schweizerischer Bankpersonalverband“, “Doctors’ associations“, “Teachers’ association” ETC
50	Identity groups	Organisation where supporters have a selective interest in group goals (not work related)	Patients, Elderly, Students, Friendship groups, Racial or ethnic, Gender “ProSenectute”, “Procap”
60	Hobby/leisure groups	Functions in organisation for people with a common sport/leisure interest including arts and culture. This can only be coded as political if they touch on matters of public affairs.	“United Hockey Union”
70	Religious group	Organisation representing churches and religious communities	“Schweizerischer Evangelischer Kirchenbund”
80	Public interest group	Organisations aimed at the promotion of issues of general public concern (e.g., environmental protection, human rights, and consumer rights).	“WWF”, “ICRC”
90	Other interest group

Table 14: Specification of political function secondary code: interest group codes

6.1.4 Secondary Political Function Coding: Bi-Cameral System (Lower House, Upper House) Applies to: X LE:

1. With the secondary codes “LH” meaning lower house or “UH” meaning upper house, we distinguish the chambers in a bi-cameral system. With this secondary code it is possible to code representatives of the specific chamber. If there is no bi-cameral system on a certain level, it will be interpreted as a higher level and therefore coded on LE-level. We will only code on LE-level when there is no bi-cameral system.
2. Committees which have members in both chambers will also be coded on the general LE-level.

6.1.5 Political functions coding: some final specifications

1. Writing numbers: Sometimes, there are more than 10 categories. To keep the hierarchy intact, numbers should always be written as 1 to 9 as 01 to 09 (at least for levels 4).
2. Delimiters: A code consisting of numbers and letters could be difficult to read. Especially if differences in length appear due to more information (policy area). Therefore, we introduce delimiters in the code. A possible delimiter might be “_”. While not strictly necessary, delimiters would facilitate reading the codes and keeping an overview. Ex. The code for the Chair of the EP’s standing committee on Internal Market and Consumer Protection B.A_02.01_01 (where “B” denotes the EU level, “A” the legislative, “02” Committee / delegation / faction, the first “01” the policy area, and second “01” means leadership)
3. Missingness: If information is missing on policy area, then this will be indicated by “XX”. Ex. The code for a municipal councillor might be: I.B.1_XX.03.

7 Parliament Data Frame (PARL)

The parliament data frame (PARL) includes variables that contain relevant information related to the parliament primarily as an institution. It aims to cover both stable and time-variant characteristics of national and regional parliaments. The characteristics of parliaments change over time, differ between countries and differ between levels within countries. A “parliament” thus always entails a combination of a legislative term (e.g. `leg_period`) and either a country (`country_abb`) or a within country region (`region_abb`).

Table 15: Summary of Parliament Level Variables (PARL)

NAME	TYPE	VALUES/EXAMPLE ¹³	SHORT DESCRIPTION	EFFORT
<code>parliament_id</code>	String	[country_abb]_[level] -[type/reg_abb]_[year] NL_NT-TK_1946 DE_NT-BT_2009 CH_NT-NR_1971 CH_RE-BS_1997 DE_RE-BW_1956	parliament identifier: Unique code for different parliaments, both across countries/levels and over time.	COMP
<code>leg_period.start</code>	Date	29apr1976 aug1953 2012	legislative period start date: Variable that captures when a specific legislative period started.	COMP
<code>leg_period.end</code>	Date	29apr1976 aug1953 2012	legislative period end date: Variable that captures when a specific legislative period ended.	COMP
<code>level</code>	String	NT RE	Location in the multi-level structure: The organisational level at which the parliament is situated	COMP
<code>country_abb</code>	String	CH = Switzerland DE = Germany NL = Netherlands	country abbreviation of the parliament	COMP
<code>region_abb</code>	String	BB = Brandenburg ZH = Zürich	abbreviation of the region: abbreviation of the name of the region (NL), federal state (DE) or canton (CH) of the parliament	COMP
<code>leg_period</code>	Integer	43 13	number of the legislative period	CWA ¹⁵
<code>election_date</code>	Date	18oct1987	date of election	COMP
<code>assembly_name</code>	String	Kantonsrat Grosser Rat Buergerversammlung Landesparlament	name of assembly: as locally used	COMP
<code>assembly_abb</code>	String	Bundestag(BT) Tweede Kamer(TK) Kantonsrat (KR) Grosser Rat (GR) Landesparlament (LP)	two letter abbreviation of the assembly	COMP
<code>election_type</code>	String	regular early	Variable indicating if the election was a normal or an early election due to special circumstance (like a crisis and constitutional change)	
<code>coalition_parties</code>	String	NL_VVD_NT; NL_PvdA_NT	Array of party_id's indicating which part(ie /y) was / where a part of the governing coalition during this parliamentary term	

¹⁵For regional parliaments that have existed since the middle ages, counting legislative periods might not be feasible.

Table 15: Summary of Parliament Variables (Continued)

NAME	TYPE	VALUES/EXAMPLE ¹⁴	SHORT DESCRIPTION	EFFORT
<code>previous_parliament</code>	ID	NL_NT-TK_1946	Parliament_id of the previous parliament	

parliament_id: Unique identification code for different parliaments, both across countries/level and over time. The id consists of the country abbreviation (variable `country_abb`), the organizational level of the parliament (see `level`), the type (DE: BundesTag(BT) & BundesRat(BR); NL: Tweede Kamer(TK) & Eerste Kamer(EK); CH: National Rat (NR) & StandeRat (SR), the abbreviated region (variable `reg_abb`), and the start year of the legislative period (see variable “`leg_period_start`”) appended by underscore.

leg_period_start: Contains the official start date, in the same format as described in variable “`birth_date`” (two digit day, month (first three letters), and four digit year). The chosen date is the official date mentioned by the parliament as the first day of the respective legislative term, if available. Otherwise the day after the election is used.

leg_period_end: Contains the official end date, in the same format as described in variable “`birth_date`” (two digit day, month (first three letters), and four digit year). The chosen date is the official date mentioned by the parliament as the last day of the respective legislative term, if available, otherwise the day before the next election is used.

country_abb: Abbreviated name of the country the parliament (parliamentarian) belongs to (works for). The used abbreviations follow the ISO “ALPHA-2 code. (See: http://www.nationsonline.org/oneworld/country_code_list.htm)

region_abb: Abbreviation of the name of the region (NL), federal state (DE), or canton (CH) of the parliament using two capital letters.

election_date: Contains the official date(s) the respective parliament was elected, in the same format as described in variable “`birth_date`” (two digit day, month (first three letters), and four digit year).

assembly_name: String variables that captures the assembly name as locally used. For the writing of the assembly name, refer to the rules explained regarding the variable `last_name`.

8 Membership Episode Data Frame (MEME)

The membership episode data frame (MEME) includes variables that relate to instances of membership in parties (both national and regional) of an individual. The characteristics of member episodes change over time and differ between individuals. A “member episode” entails a combination of the party in question (i.e. `party_id`), an id specific to the individual politician (`pers_id`), and the id of the party (`part_id`). An example of this data-frame can be downloaded at: [URL].

Table 16: Summary of Membership Episode Variables (MEME)

NAME	TYPE	VALUES/EXAMPLE ¹⁶	SHORT DESCRIPTION	EFFORT
<code>memep_id</code>	String	CH_Abate.Fabio.1966 _CH_RL_NA__2	membership episode identification code: a combination of <code>pers_id</code> , <code>party_id</code> and an integer to count episodes.	COMP
<code>pers_id</code>	String	CH_Abate Fabio.1966 DE_Mueller Johann.1956.dec	parliamentarians’ individual identification code: a combination of <code>country_abb</code> , the <code>last_name</code> , the <code>first_name</code> and <code>birth_date</code>	COMP
<code>party_id</code>	String	DE_CSU_RE-BY	unique party identifier: a combination of <code>country_abb</code> , <code>party_abb</code> , <code>level</code> , <code>reg_abb</code>	COMP

Table 16: Summary of Membership Episode Variables (MEME) (Continued)

NAME	TYPE	VALUES/EXAMPLE ¹⁷	SHORT DESCRIPTION	EFFORT
<code>memep_startdate</code>	Date	29apr1976 aug1953 2012	membership episode start date: Variable that captures when a specific membership episode started.	COMP
<code>memep_enddate</code>	Date	29apr1976 aug1953 2012	membership episode end date: Variable that captures when a specific membership episode ended.	COMP
<code>memep_type_raw</code>	String	honorary member	membership episode type (raw): a string variable that describes the kind of party membership the individual in question has. ¹⁸	CWA

memep_id: Identification code that is specific to this data frame. It identifies an individual (`per_id`) who is a member of a specific party (`party_id`).

pers_id personal identifier: Identification code, individual to each parliamentary ACROSS levels. The code consists of the country abbreviation (as specified in variable `country_abb`), the last name (as specified in variable `last_name`). The first name (as specified in `first_name`) and the year of birth (as specified by the last four digits of variable `birth_date`) connected by underscores. In case this id is not unique to one person, the month of birth (as specified in the middle of the `birth_date` variable string) is appended by underscore. If the month of birth is not known or does still not uniquely identify the person, the party abbreviation (as specified by the variable `party_abb`) is appended (additionally to the month of birth) by underscore.

Alternative IDs and how they correspond to the main ID can be found in the appendix.

party_id unique party identifier: party identifier consisting of the country abbreviation (see 7, variable `country_abb`), the party abbreviation (see `party_abb`), the organizational level and the abbreviated region (see 7, variables `level`, and `reg_abb`) appended by underscore.

memep_startdate membership episode start date: Contains the start date of the membership episode, in the same format as described in variable `birth_date` (two-digit day, month (first three letters), and four-digit year). The chosen date is either: i) the date on which the MP in question took up the post; ii) the official date mentioned by the parliament as the first day of the respective legislative term to which the membership episode corresponds.

memep_enddate: membership episode end date Contains the end date of the membership episode, in the same format as described in variable `birth_date` (two digit day, month (first three letters), and four digit year). The chosen date is either: i) the date on which the MP in question left the post; ii) the official date mentioned by the parliament as the last day of the respective legislative term to which the membership episode corresponds.

9 Party Data Frame (PART)

The party data frame (PART) includes variables that relate to the parties included in the dataset. The characteristics of parties change over time, between countries, and across regions. A “party” entails a combination of the country (`country_abb`), the party (`party_abb`), the level at which the party operates (`level`), and the region (`reg_abb`). An example of this data-frame can be downloaded at: [URL].

¹⁸If no information, defaults to “NA”.

Table 17: Summary of Party Variables (PART)

NAME	TYPE	VALUES/EXAMPLE ¹⁹	SHORT DESCRIPTION	EFFORT
party_id	String	[country_abb]- [party_abb]- [level] [reg_abb] “DE.CSU.RE- BY”	unique party identifier: A variable that uniquely identifies each party included in the dataset. Combination of <code>country_abb</code> , <code>party_abb</code> , <code>level</code> , and <code>reg_abb</code> , where each element is separated by an underscore, and the regional abbreviation by a hyphen.	COMP
party_level	String	see level 2! (6.1.2)	organizational/federal level a party operates on	COMP
region_abb	String	<ul style="list-style-type: none"> • BB = Branden- burg • ZH = Zürich 	abbreviation of the region: abbreviation of the name of the region (NL), federal state (DE) or canton (CH) of the parliament/district	COMP
ancestor_party_id	String	NL.ARP.NA	ancestor party id: shows the party id of other party(s) if a party has came about from a merger or after a split from another party	CWA
mother_party_id	String	NL.CDA.NA	mother party id: gives the id of (the) mother party (if applicable).	CWA
party_abb	String	<ul style="list-style-type: none"> • CDA = Dutch Christian Democrats • CDU = Ger- man Christian Democrats 	abbreviated party names: abbreviated party names as used by ParlGov	COMP
party_name	String	<ul style="list-style-type: none"> • Buendnis 90/ Die Gruenen 	party names: Full party name in original language.	CWA
party_parlgov_id	Integer	<ul style="list-style-type: none"> • 1591 • 1277 	ParlGov party id: party id as used by ParlGov	CWA
old_party_abb	Integer	<ul style="list-style-type: none"> • KK-CC;CCPS;K- CPS 	Old party id: a list of party abbreviations that the party has had in the past, separated by semi-colons	CWA ²¹

²¹Only collected for CH in the interest of identifying party affiliations for members.

Table 17: Summary of Party Variables (PART) (Continued)

NAME	TYPE	VALUES/EXAMPLE ²⁰	SHORT DESCRIPTION	EFFORT
old_party_name	Integer	Catholic Conservative; Katholische Konservative; Conservateurs catholiques/Conservative-Christian-Social People's Party; Katholisch-Konservativ-Christlichsoziale Volkspartei; Parti Conservateur Chretien Social/Conservative-Christian-Social People's Party; Katholisch-Konservativ-Christlichsoziale Volkspartei; Parti Conservateur Chretien Social	Old party name: a list of party names that the party has had in the past, separated by semi-colons	CWA ²²

party_id unique party identifier: party identifier consisting of the country abbreviation (see 7, variable `country_abb`), the party abbreviation (see `party_abb`), the organizational level and the abbreviated region (7, variables `level`, and `reg_abb`) appended by underscore.

party_level organizational/federal level a party operates on.

region_abb region abb abbreviation region: abbreviation of the name of the region (NL), federal state (DE), or canton (CH) of the parliament using two capital letters.

ancestor_party_id ancestor party id: shows the party id of other party(s) if a party has came about from a merger or after a split from another party (if applicable). Can be empty.

mother_party_id mother party id: gives the id of (the) mother party (if applicable). Signifies relations between parties in the multilevel structure (e.g. but not limited to lower level membership implying higher level membership). Can be empty.

party_abb abbreviated party names: party names as used by ParlGov, consisting of a party abbreviation appended to a three digit country abbreviation by underscore.

party_name: party names Full party name in original language: the same spelling rules as set out for the variable "last name" apply. Furthermore the rules set out by ParlGov apply. See: <http://www.parlgov.org/documentation/codebook>.

party_parlgov_id party id: party id as used by ParlGov (see <http://www.parlgov.org/documentation/codebook>).

old_party_abb old party abb: a(n array of) party potential abbreviation(s) that the party has had in the past, separated by semi-colons.

old_party_name old party name: a(n array of) potential party name(s) that the party has had in the past, separated by semi-colons.

²²Only collected for CH in the interest of identifying party affiliations for members.

10 Faction Data Frame (FACT)

The faction data frame (FACT) includes variables that relate to groups of parties in assemblies (i.e. 'factions', or 'parliamentary party groups'). The characteristics of factions change over time, across parliaments, and within parliaments. We define a 'faction' as collection of individuals in parliament that collectively represent either a party or a collection of parties. Not that this thus entails both the group made up entirely by members of the same party, like is for example the case in the Netherlands, as party groups like is typical in Germany and Switzerland. Because characteristics of party groups can differ from the characteristics of their subparties (e.g. number of seats, gender quotas), in later cases both the parties collectively and and their party specific subgroups will occur in the data. The DE_NT_2009__ DE_CDU_NT-DE_CSU_NT group will as such occur three times in the data. Once as "DE_NT_2009__ DE_CDU_NT-DE_CSU_NT", but also as "DE_NT_2009__DE_CDU_NT" and as "DE_NT_2009__DE_CSU_NT" **(We should discuss this as it adds a lot of redundant information: Wouldn't it make more sense to connect QUOT to PART and add start and end dates for quotas in QUOT? Then, there would be no redundant entries in FACT. For seats per party within the same faction, faction composition already covers this.)**

Table 18: Summary of Faction Variables (FACT)

NAME	TYPE	VALUES/EXAMPLE ²³	SHORT DESCRIPTION	EFFORT
<code>faction_id</code>	String	DE_NT_2009__ DE_CDU_NT- DE_CSU_NT CH_NT- NR_2009__CH_BDP__1	faction identifier: combination of the parliament_id, and the party ids in alphabetical order of the party(ies) that make(s) up the faction,seperated by double hypens '-'. Changes in the same faction within the same year are denoted by a count after a double underscore '__'.	COMP
<code>faction_name</code>	String	<ul style="list-style-type: none"> • gruene-fraktion • union 	faction name: the name of the faction.	COMP
<code>party_id('s)</code>	String	DE_CDU_NT; DE_CSU_NT	party identifiers for faction members: (array of) party identifiers for all factions included in the faction, separated by a semi-colon.	COMP
<code>parliament_id</code>	String	[country_abb]_[level] -[type/reg_abb]_[year] NL_NT-TK_1946 DE_NT-BT_2009 CH_NT-NR_1971 CH_RE-BS_1997 DE_RE-BW_1956	parliament identifier: Unique code for different parliaments, both across countries/levels and over time.	COMP
<code>seats</code>	Integer	<ul style="list-style-type: none"> • 45 • 27 	Faction seats: the number of seats that all parties included in the faction have combined in the parliament in question.	COMP
<code>faction_composition</code>	Integer	<ul style="list-style-type: none"> • 29-16 • 20-7 	Faction composition: the number of seats that per party in the faction, separated by hyphen. Order follows the order of abbreviations in <code>faction_id</code>	COMP

Table 18: Summary of Faction Variables (FACT) (Continued)

NAME	TYPE	VALUES/EXAMPLE ²⁴	SHORT DESCRIPTION	EFFORT
<code>faction_start</code>	Date	29apr1976 aug1953 2012	faction start date: Variable that captures when a specific faction came into being or changed with regard to any other collected faction variable.	COMP
<code>faction_end</code>	Date	29apr1976 aug1953 2012	faction end date: Variable that captures when a specific faction ended or changed with regard to any other collected faction variable.	COMP

faction_id **faction identifier:** combination of the name of the faction (`faction_name`) and `parliament_id`.

faction_name **name of the faction:** mentions the full name of the faction (as given by the relevant parliament/in the original language). As for spelling of the names, see rules at `last_name_raw`.

party_id('s) **unique identifiers for faction parties:** party identifiers for all parties included in the faction; each party identifier consists of the country abbreviation (see 7, variable `country_abb`), the party abbreviation (see `party_abb`), the organizational level and the abbreviated region (7, variables `level`, and `reg_abb`) appended by underscore.

parliament_id: Unique identification code for different parliaments, both across countries/level and over time. The id consists of the country abbreviation (variable `country_abb`), the organizational level of the parliament (see `level`), the type (DE: BundesTag(BT) & BundesRat(BR); NL: Tweede Kamer(TK) & Eerste Kamer(EK); CH: National Rat (NR) & StandeRat (SR), the abbreviated region (variable `reg_abb`), and the start year of the legislative period (see variable “`leg_period_start`”) appended by underscore.

seats: the number of seats that the faction has in the respective parliament.

fact_composition **composition of the faction:** the seats held by parties that constitute the faction, in the same order as the parties contained in `faction_id`.

faction_start: This variable gives the start date of a faction. These date entries consist of the two-digit day, the month (first three letters, small, of the English name of the month) and the four-digit year. In case of missing information on days or month, shortened versions of the date are used. In cases when the sources do not provide episode data, periods are inferred based on snapshots of faction compositions in conjunction with parliamentary episodes.

faction_end: This variable gives the end date of a faction. These date entries consist of the two-digit day, the month (first three letters, small, of the English name of the month) and the four-digit year. In case of missing information on days or month, shortened versions of the date are used. In cases when the sources do not provide episode data, periods are inferred based on snapshots of faction compositions in conjunction with parliamentary episodes.

11 Committee Data Frame (COMM)

The committee data frame (COMM) includes variables that relate to non-partisan specific groups in legislative assemblies (i.e. ‘parliamentary committees’, ‘parliamentary delegations’, ‘all-party parliamentary groups’, and their respective subbodies). These bodies can either have official standing in parliament such as committees and delegations or constitute informal cross-party groups such as all-party parliamentary groups. In either case, the characteristics of these legislative bodies change over time, across parliaments, and within parliaments.

Table 19: Summary of Faction Variables (COMM)

NAME	TYPE	VALUES/EXAMPLE ²⁵	SHORT DESCRIPTION	EFFORT
comm_id	PriID	CH_NT-NR_1999__ CO FK-SC EDA/WBF CH_NT-NR_2003__CH_ NT-SR_2003__DL Del-FL CH_NT-NR_2011__CH_ NT-SR_2011__AG Wohn - und Grundeigentum	committee identifier: combination of the parliament_id(s), the committee maintype, the abbreviated maintype name, as well as the subtype and the abbreviated subtype name if applicable. parliament_id and maintype are separated by double underscores '__'. Maintype/subtype and abbreviated maintype/subtype name are separated by a vertical bar ' '. In case a subtype is specified, it is separated from the abbreviated maintype name by double hypens '- '.	COMP
parliament_id	ID	NL_NT-TK_1946 DE_NT-BT_2009 CH_NT-NR_1971 CH_RE-BS_1997 DE_RE-BW_1956	parliament identifier: Unique code for different parliaments, both across countries/levels and over time.	COMP
comm_name	String	Delegation bei der parlamentarischen Versammlung der OSZE Geschäftsprüfungs- kommission Parlamentarische Gruppe Kultur	committee name: name of the committee, delegation, all-party parliamentary group, or subcommittee.	COMP
comm_maintype	String	CO DL AG	type of committee: Body maintype is either parliamentary committee (CO), parliamentary delegation (DL), or all-party parliamentary group (AG). Maintype uses two-letter abbreviations.	COMP
comm_subtype	String	SC SD SA	subtype of committee: Body subtype refers to either parliamentary subcommittee (SC), parliamentary subdelegation (SD), or all-party parliamentary subgroup (SA). Subtype uses two-letter abbreviations.	CWA
comm_maintype_abb	String	FK FinDel Drogenpolitik	main committee abbreviation: Abbreviation of main body as used by officials. If none is available, a shortened version of comm_name is used instead.	COMP
comm_subtype_abb	String	EDA/WBF AgrRisikorep D	subcommittee abbreviation: Abbreviation of subbody as used by officials. If none is available, a shortened version of comm_name is used instead.	CWA

Table 19: Summary of Committee Variables (COMM) (Continued)

NAME	TYPE	VALUES/EXAMPLE ²⁶	SHORT DESCRIPTION	EFFORT
seats	Integer	<ul style="list-style-type: none"> • 45 • 27 	committee seats: the number of seats that all parties included in the respective body have combined during the respective parliament.	CWA
comm_start	Date	29apr1976 aug1953 2012	committee start date: Variable that captures when a specific committee, delegation, or all-party parliamentary group came into being or changed with regard to any other collected COMM variable.	COMP
comm_end	Date	29apr1976 aug1953 2012	committee end date: Variable that captures when a specific committee, delegation, or all-party parliamentary group ended or changed with regard to any other collected COMM variable.	COMP

comm_id: Combination of the parliament_id(s), the committee maintype, the abbreviated maintype name, as well as the subtype and the abbreviated subtype name if applicable:

[parliament_id(s)]_--[comm_maintype][[comm_maintype_abb]--[comm_subtype][[comm_subtype_abb]

Multiple parliament_id as well as the last parliament_id and the maintype are separated by double underscores '_'. Maintype/subtype and abbreviated maintype/subtype name are separated by a vertical bar '|'. In case a subtype is specified, it is separated from the abbreviated maintype name by double hypens '-'

parliament_id: Unique identification code for different parliaments, both across countries/level and over time. The id consists of the country abbreviation (variable **country_abb**), the organizational level of the parliament (see **level**), the type (DE: Bundestag(BT) & Bundesrat(BR); NL: Tweede Kamer(TK) & Eerste Kamer (EK); CH: Nationalrat (NR) & Staenderat (SR), the abbreviated region (variable **reg_abb**), and the start year of the legislative period (see variable "leg_period_start") appended by underscore.

comm_name: Official long version of the name of the committee, delegation, all-party parliamentary group, or subcommittee.

comm_maintype: Main type refers to either parliamentary committee (CO), parliamentary delegation (DL), or all-party parliamentary group (AG). Main type uses two-letter abbreviations.

comm_subtype: Subtype refers to either parliamentary subcommittee (SC), parliamentary subdelegation (SD), or all-party parliamentary subgroup (SA). Subtype uses two-letter abbreviations.

comm_maintype_abb: Abbreviation of main type as used by officials. If none is available, a shortened version of comm_name is used instead.

comm_subtype_abb: Abbreviation of subtype as used by officials. If none is available, a shortened version of comm_name is used instead.

seats: The number of seats that all parties included in the respective body have combined during the respective parliament.

comm_start: Variable that captures when a specific committee, delegation, or all-party parliamentary group came into being or changed with regard to any other collected COMM variable. These date entries consist of the two-digit day, the month (first three letters, small, of the English name of the month) and the four-digit year. In case of missing information on days or month, shortened versions of the date are used. In cases when the sources do not provide episode data, periods are inferred based on snapshots of faction compositions in conjunction with parliamentary episodes.

comm_end: Variable that captures when a specific committee, delegation, or all-party parliamentary group ended or changed with regard to any other collected COMM variable. These date entries consist of the two-digit day, the month (first three letters, small, of the English name of the month) and the four-digit year. In case of missing information on days or month, shortened versions of the date are used. In cases when the sources do not provide episode data, periods are inferred based on snapshots of faction compositions in conjunction with parliamentary episodes.

12 Electoral List Data Frame (ELLI)

The electoral list data frame (ELLI) includes information on the individual electoral lists, which might originate from one or several parties. All variables in this data frame are on the level of the electoral lists and vary per parliamentary term: TV_P (start/end 'parliament_id'). Each list is identified by a combination of the parliament id (parliament_id), the district id (district_id), and the name of the list(list_name). An example of this data-frame can be downloaded at: http://parliamentarycareersincomparison.org/wp-content/uploads/2017/03/PCC-example-data_v311.xlsx

Table 20: Summary of the Electoral List Level Variables (ELLI)

NAME	TYPE	VALUES/EXAMPLE	SHORT DESCRIPTION	EFFORT
list_id	PriID	[parliament_id]--[district_id] --[list_name] DE_NT_2009-- CH_NT_1971 --Zuerich --Landesliste- Sozialdemokratische- Partei-Deutschlands	unique list identifier: consisting of the parliament id, the list name and the district id	COMP
list_name	String	<ul style="list-style-type: none"> • Sozialdemokratische- und- Gewerkschaftliche- Liste • NA_ZH.1 	name of the electoral list: mentions the full name of the electoral list (as given by the relevant electoral office) ²⁷ . For CH only: In case the MP did not run on a list, we use the format NA_[kanton_abb]_index, where index starts at 1 and increments by 1 with every unnamed list for the Kanton in the order in which they are listed in the <i>Bundesblatt</i>	COMP
parliament_id	ID	[country_abb]--[level] --[type/reg_abb]--[year]	parliament identifier: Unique code for different parliaments. See parliaments date-frame on page 7	COMP
party_id	ID	[country_abb]--[party_abb] --[level]--[reg_abb] DE_CSU_RE-BY	unique party identifier (see above)	COMP
district_id	ID	[parliament_id] --[constituency_name] CH_NT_1971 --Zuerich	unique district identifier: identification code of the district, as used in the other data frames, identifying the district the electoral list belongs to.	COMP

²⁷In case a full list name (string does not exist, we use the list number (list_number)

list_level	String	<ul style="list-style-type: none"> • DE.BY²⁸ – list on the German level of Bavaria • CH.GR – list on the level of the Swiss Canton Graubunden 	level a list is on: the variable specifies the level (national or subnational) a list is on and gives the identity of the level (e.g. Germany federal, or German State of Bavaria)	COMP
list_number	Numeric	<ul style="list-style-type: none"> • 1 • 13 	number of electoral list: identifying number of consecutive number of the list as mentioned in the original data source (usually the national bureau for elections)	CWA
list_party_id	String (of IDs)	[country_abb]-[party_abb]-[level]-[reg_abb] DE.CSU.RE-BY	list party id(s): id of the party(s) that constructed the list. Can be identical with the variable “party_id”. Several party ids can exist in this variable, if the list is a collaboration of several parties	COMP
list_length	Numeric	<ul style="list-style-type: none"> • 1:99 	total number of spots on the electoral list	CWA
list_connections	String (of IDs)	CH.NA.Sozialdemokratische-und-Gewerkschaftliche-Liste_36	id(s) of list(s) an electoral list is connected with	CWA
list_votes_total_source	Numeric	<ul style="list-style-type: none"> • 1250 	total number of valid votes a list got	CWA

list_id unique list identifier: electoral list identifier consisting of the parliament id (see variable “parliament_id” in 7, consisting of variables [country_abb]-[level]-[reg_abb]-[leg_period]), the list name (see “list_name”), the identity of the list level (see “list_level”) connected by underscores.

list_name name of the electoral list: mentions the full name of the electoral list (as given by the relevant electoral office). As for spelling of the names, see rules at “last_name_raw”. Blanc spaces between the words are exchanged by hyphens. When a list_name is unknown NA_[reg_abb]-[integer] is used.

parliament_id Unique identification code for different parliaments. See PARL for details on page 7.

party_id unique party identifier: party identifier. See PART for details on page 9

list_level level the list is on: the variable consists of the country abbreviation (two letters) and an abbreviation of the sub-national unit (region/state/or sub- and inter-regional units in the case of the NL, again two letter abbreviations) appended to each other by underscore. The variable specifies the level (national or subnational) a list is on (a national party list would have a country abbreviation but no sub-national/regional abbreviation) and gives the identity of the level (e.g. DE_XX = German federal level list, or DE_BY = state based list in the German state of Bavaria).²⁹

²⁸In some cases (e.g. the Netherlands) a list of electoral district abbreviations will be used instead of regional abbreviations

²⁹Note that the variable “list_level” differs from the variable **level**, in that it does not only indicate the organizational level (e.g. national or regional), but also the identity of the specific level.

list_number **number of electoral list:** identifying number of consecutive number of the list as mentioned in the original data source (usually the national bureau for elections).

list_party_id **list party id(s):** id of the party(s) that constructed the list. Can be identical with the variable “party_id”. Several party ids can exist in this variable, if the list is a collaboration of several parties. on the construction of party ids see ?? variable “party_id”.

list_length **total number of spots on the electoral list**

list_connections **list_connections:** this variable is NL - specific and mentions the list_id(s) of other lists that the respective list is connected with, which means that these lists will get the votes that the respective list can not use (e.g. because it did not make the quorum).

list_votes_total_source **total number of valid votes a list got** according to a specific source.

13 Electoral List Entry Data Frame (ELEN)

The electoral list entry data frame (ELEN) includes information on the candidacy of individual politicians (on lists). All variables in this data frame are on the level of the individual politician (see **pers_id**) and vary per election (see **parliament_id**) and election district see **list_id**. Each election list entry is identified by a combination of the electoral list id (**list_id**), and the ID identifying the person (**pers_id**). An example of this data-frame can be downloaded [HERE](#)

Table 21: Summary of the Electoral List Entry Variables (ELEN)

NAME	TYPE	VALUES/EXAMPLE	SHORT DESCRIPTION	EFFORT
elec_entry_id	PriID	[list_id]__[pers_id] DE_NT_2009 __DE_NT_2009 __Saarland __Landesliste- Sozialdemokratische- Partei-Deutschlands __DE_Funk __Alexander_1974	unique electoral list entry identifier: consisting of the the electoral list id (list_id), and the personal id (pers_id) ³⁰	COMP
list_id	ID	[parliament_id]__[disctrict_id] __[list_name] DE_NT_2009__DE_NT_2009 __Saarland __Landesliste- Sozialdemokratische- Partei-Deutschlands	unique list identifier: consisting of the parliament id, the list name and the list level	COMP
pers_id ³¹	ID	CH_Abate __Fabio_1966 DE_Mueller __Johann_1956_dec	parliamentarians individual identification code: a combination of country_abb , the last_name , the “first_name” and the birth_date	COMP
listplace	Numeric	1:99	listplace the person ran on	COMP

³⁰In case the same person appears more than once on the same electoral list, a consecutive number is appended by underscore (e.g.: DE_NT_2009 __DE_NT_2009 __Saarland __Landesliste-Sozialdemokratische-Partei-Deutschlands __DE_Funk __Alexander_1974_2)

³¹To ensure compatibility of datasets and make usage and merging of our dataset easier we will add several id variables. The first one (**pers_id**) will be constructed by us, consists of the country, last name, first name, and birth year (and possibly a running number), and is unique to parliamentarians across countries and levels. This ‘information based’ ID format allows consistent IDS to be constructed across a variety of sources while minimizing the need to look-up IDS in the already collected data.

elected_ [source]	Binary	0=not elected 1=elected	elected_[source]: binary variable that specifies a list position or other type of candidature of a candidate was successful (elected) (according to a specific source (mentioned in [external data source]))	CWA
candidature_type	String	L = list D = district LD = list and district	type of candidature of a person: only relevant for DE	CWA
seat_type	Numeric	1 direct mandate 2 list without succession 3 succession via list 11 byelection 12 Volkskammer GDR 13 list	type of seat/mandate in detail: only relevant for DE	CWA
candidate_votes	Numeric	1250	total number of votes a candidate got in the elections for the national parliament	CWA

elec_entry_id unique electoral list entry identifier: consisting of the the electoral list id (**list_id**), consisting of [parliament_id]_[district_id]_[list_name], and the personal id (**pers_id**), which is a combination of **country_abb**, the **last_name**, the “first_name” and the **birth_date**, connected by underscores.

list_id unique list identifier: electoral list identifier consisting of the parliament id (see variable “parliament_id” in 7, consisting of variables [country_abb]_[level]_[reg_abb]_[leg_period]), the list name (see “list_name”), the identity of the list level (see “list_level”) connected by underscores.

pers_id parliamentarians individual identification code: Identification code, individual to each parliamentarian ACROSS levels. The code consists of the country abbreviation (as specified in variable **country_abb**), the last name (as specified in variable **last_name**). The first name (as specified in **first_name**) and the year of birth (as specified by the last four digits of variable **birth_date**) connected by underscores. In case this id is not unique to one person, the month of birth (as specified in the middle of the **birth_date** variable string) is appended by underscore. If the month of birth is not known or does still not uniquely identify the person, the party abbreviation (as specified by the variable **party_abb**) is appended (additionally to the month of birth) by underscore.

Alternative IDs and how they correspond to the main ID can be found in the appendix.

listplace Variable is a positive number that indicates the position on the electoral list a person ran on and/or was elected on.

candidature_type Variable is mainly important for German candidate selection and mentions the type of candidature (list, district, or both) of a candidate.

seat_type Variable is mainly important for (elected) parliamentarians from mixed systems like Germany. It specifies the seat type or mandate of a parliamentarian. The parliamentarian can either have a direct mandate (based on a district mandate, 1), can be part of the parliament due to a successful list candidacy without succession (2), he/she could have succeeded another parliamentarian via the list (3), could have entered the parliament due to a byelection (11), could have been part of the GDR Volkskammer (12), or can be part of the parliament due to a successful list candidacy (13).

candidate_votes Numeric variable mentioning the total number of (valid) votes a candidate/list position got in (national parliament) elections.

elected_[source]: Binary variable that specifies whether, according to a specific source (mentioned in [external data source]), the federal statistical bureau for example, a list position or other type of candidature of a candidate was successful (elected) or not.

14 Electoral District Data Frame (ELDI)

The electoral district data frame (ELDI) includes information on the organizational units (districts) the elections are executed in. For example, the name of a given district or size of it. Variables in this data frame are on the level of the districts and (are allowed to) vary per election (see ‘parliament_id’). Each electoral district (constituency) is identified by a combination of the parliament (parliament_id) and the constituency name (constituency_name). An example of this data-frame can be downloaded [HERE](#)

Table 22: Summary of District Level Variables (ELDI)

NAME	TYPE	VALUES/EXAMPLE	SHORT DESCRIPTION	EFFORT
district_id	PriID	[parliament_id] --[constituency_name] CH_NT_1971 --Zuerich	unique district identifier: identification code as used in the other data frames	COMP
parliament_id	ID	[country_abb]-[level]-[year] NL_NT_1946 DE_NT_2009 CH_NT_1971 [country_abb]-[level] -[reg_abb]-[year] CH_RE-BS_1997 DE_RE-BW_1956	parliament identifier: Unique code for different parliaments, both across countries/levels and over time.	COMP
country_abb	String	<ul style="list-style-type: none"> • CH = Switzerland • DE = Germany • NL = Netherlands 	country abbreviation of the parliament	COMP
region_abb	String	<ul style="list-style-type: none"> • BB = Brandenburg • ZH = Zürich 	abbreviation of the region: abbreviation of the name of the region (NL), federal state (DE) or canton (CH) of the parliament/district	COMP
constituency_name	String	<ul style="list-style-type: none"> • Zuerich • Stuttgart 	name of the constituency	COMP
constituency_code_CLEA	Numeric	<ul style="list-style-type: none"> • 001–900 • 901–999 	identifying code of the constituency: as assigned by CLEA	CWA
constituency_name_CLEA	Numeric	<ul style="list-style-type: none"> • 001–900 • 901–999 	name of the constituency: as assigned by CLEA	CWA

dist_magnitude	Numeric	<ul style="list-style-type: none"> • -992 = Uncontested election (i.e. a single candidate contested the election) • -994 = Suspended election 	district magnitude: measured by the number of seats allocated in a given district	CWA
district_votes_total_source	Numeric	<ul style="list-style-type: none"> • 1250 	total number of valid votes: cast within a given district and period	CWA

district_id **unique district identifier** unique identification code for electoral districts consisting of the parliament id (variable “parliament_id”), and the name of the constituency (variable “constituency_name”) appended by underscore.

parliament_id Unique identification code for different parliaments. See PARL for details.

country_abb **country abbreviation of the parliament:** abbreviated name of the country the parliament/district belongs to. The used abbreviations follow the ISO “ALPHA-2 code. (See: http://www.nationsonline.org/oneworld/country_code_list.htm)

region_abb **abbreviation region:** abbreviation of the name of the region (NL), federal state (DE), or canton (CH) of the parliament using two capital letters. For CH Based on [urlhttp://swiss-government-politics.all-about-switzerland.info/swiss-federal-states-cantons.html](http://swiss-government-politics.all-about-switzerland.info/swiss-federal-states-cantons.html)

district_aliases **district_aliases** (array of) potential aliases the district it also known under. E.g. ‘kies kring numbers’ for NL.

constituency_name **name of the constituency:** mentions the full name of the electoral constituency (as given by the relevant electoral office/in the original language). As for spelling of the names, see rules at “last_name_raw”. Called ‘CST’ in CLEA. Dots are removed. Remaining spaces are replaced by hypens (‘-’)

constituency_code_CLEA **identifying code of the constituency:** As assigned by the Constituency Elections Archive (CLEA) (see <http://www.electiondataarchive.org/about.html>). Note: this code is unique only to constituency and election (hence, the same district can have different codes at different elections). CLEA codes exist only for some national elections.

dist_magnitude **district magnitude:** measured by the number of seats allocated in a given district. The number of seats is expressed in a positive number, the following other codes are possible:

- -992 = Uncontested election (i.e. a single candidate contested the election)
- -994 = Suspended election

(See also [urlhttp://www.electiondataarchive.org/about.html](http://www.electiondataarchive.org/about.html))

district_votes_total_[source] **total number of valid votes:** a list got within a given district and period according to a given data source (which is mentioned in the end of the variable name).

15 Organisations Module(ORGS)

The organisations data frame (ORGS) captures information on the level of organisations that parliamentarians are members of at some stage of their career. In most instances, these organisations are interest groups as data on parties (PART), factions (FACT), and parliaments (PARL) are collected in other data frames.

Collection of data on organisations is not the primary objective of this codebook. Instead of developing an elaborate coding scheme for organisations, the data presented here are drawn mostly from secondary sources. For more extensive discussions of these sources, readers are therefore referred to the codebooks and studies for which these data have been collected and coded for in the first place.

Table 23: Summary of Organisation Level Variables (ORGS)

NAME	TYPE	VALUES/EXAMPLE	SHORT DESCRIPTION	EFFORT
org_id	ID	[country_abb]-[12 digit code] CH_H8KkcewN6fGv	unique organisation identifier: identification code to link organisations mentioned in RESE	COMP
org_name	String	[name ₁]/[name ₂]/[name _n] , [place of headquarters] Cranex AG, Opfikon Crédit Suisse / Schweizerische Kreditanstalt (SKA) fenaco Genossenschaft / Fédération de coopératives agricoles, Bern	name of organisation: Name and synonyms (previously used names, names in other languages) used for an organisation separated by slashes as well as the place of the organisation's headquarters. Minimum requirement: One name.	COMP
country_abb	String	<ul style="list-style-type: none"> • CH = Switzerland • DE = Germany • NL = Netherlands 	country abbreviation of the organisation's headquarters location	COMP
legal_form	String	[legal_form ₁]; legal_form ₂ ; [legal_form _n] AG; Gen. KollG; EG	legal form: Legal form of the organisation as given in the source. Multiple legal forms are separated by semicolons.	CWA
polactive_start_date	Date	29apr1976 aug1953 2012	political activity start date: Variable that captures when an organisation was known to be politically active for the first time	CWA
polactive_end_date	Date	29apr1976 aug1953 2012	political activity end date: Variable that captures when an organisation was known to be politically active for the last time	CWA
GAVA_assoc.id	Numeric	1 329 4605	Gava association id: Id assigned to organisations by Gava et al. in the DATABASE ON INTEREST GROUPS IN THE SWISS PARLIAMENT	CWA

GAVA_assoc_name	String	Pro Natura Basel Sportstiftung Thurgau Lutecia SA Genève	Gava association name: Name of organisations as used by Gava et al. in the DATABASE ON INTEREST GROUPS IN THE SWISS PARLIAMENT	CWA
GAVA_assoc _aggregate_name	String	Pro Natura Swiss Life Industrie du bois Suisse	Gava aggregate association name: Umbrella names of organisations as used by Gava et al. in the DATABASE ON INTEREST GROUPS IN THE SWISS PARLIAMENT	CWA
GAVA_assoc _legal_form	Numeric	11 30 50	Gava legal form: Legal form of organisations as used by Gava et al. in the DATABASE ON INTEREST GROUPS IN THE SWISS PARLIAMENT	CWA
GAVA_assoc _general	String	Association Commission Societe	Gava general form of association: Form of organisations as used by Gava et al. in the DATABASE ON INTEREST GROUPS IN THE SWISS PARLIAMENT	CWA
GAVA_assoc_ig _type_name	String	Business groups Identity groups Unions	Gava interest group category: Categories of interest groups as used by Gava et al. in the DATABASE ON INTEREST GROUPS IN THE SWISS PARLIAMENT. The categories are adapted from the INTERARENA project.	CWA
GAVA_assoc _ig_code	Numeric	26 63 89	Gava interest group category code: Category codes of interest groups as used by Gava et al. in the DATABASE ON INTEREST GROUPS IN THE SWISS PARLIAMENT. The category codes are adapted from the INTERARENA project.	CWA
GAVA_assoc _ig_code	String	Private firm/company Arts Local development	Gava interest group category code name: Category code names of interest groups as used by Gava et al. in the DATABASE ON INTEREST GROUPS IN THE SWISS PARLIAMENT. The category codes are adapted from the INTERARENA project. This variable describes the numeric values of GAVA_assoc_ig_code.	CWA

GAVA_level	Numeric	1 2 4	Gava level of organisation: Level codes ranging from local to international as used by Gava et al. in the DATABASE ON INTEREST GROUPS IN THE SWISS PARLIAMENT.	CWA
GAVA_category	String	Local Regional/Cantonal International	Gava category of level: Level codes ranging from local to international as used by Gava et al. in the DATABASE ON INTEREST GROUPS IN THE SWISS PARLIAMENT. This variable describes the numeric values of GAVA_level.	CWA
GAVA_topic	Numeric	6 10 12	Gava topic: Policy area in which an organisation is active in as used by Gava et al. in the DATABASE ON INTEREST GROUPS IN THE SWISS PARLIAMENT. Based on the Comparative Agenda Project (CAP) coding scheme.	CWA
GAVA_topic _business	Numeric	1500 1505 1564	Gava topic business: For organisations that fall into the category of General Business (1500), Gava et al. provide more detailed information on the area of business in their DATABASE ON INTEREST GROUPS IN THE SWISS PARLIAMENT.	CWA
NCCR_nqid	Numeric	46 827 2608	NCCR interest group id: Id of interest groups for which data have been collected in the context of the project NCCR IP 8: MEDIATIZATION OF POLITICAL INTEREST GROUPS by Jarren et al..	CWA
NCCR_A4 Organisation	String	Alpen-Initiative Castagna, Zürich FachFrauen Umwelt	NCCR interest group name: Name of interest groups for which data have been collected in the context of the project NCCR IP 8: MEDIATIZATION OF POLITICAL INTEREST GROUPS by Jarren et al..	CWA

GK_interest_group	String	Verein Al Forno Semaine du Gout Swissgas AG	Giger & Klüber IG name: Name of interest group for which data have been collected by Giger and Klüber for their project on SWISS MPS AND THEIR INTEREST GROUP AFFILIATION.	CWA
GK_canton_ig_raw	String	LU Halle (BRD) ZH	Giger & Klüber canton of IG location: Raw (uncleaned) string of the canton in which an interest group is headquartered. These data have been collected by Giger and Klüber for their project on SWISS MPS AND THEIR INTEREST GROUP AFFILIATION.	CWA
GK_type	String	culture ngo prof	Giger & Klüber IG type: Type of interest group for which data have been collected by Giger and Klüber for their project on SWISS MPS AND THEIR INTEREST GROUP AFFILIATION.	CWA
GK_policy_code	String	7 13 28	Giger & Klüber IG policy code: Policy area of interest group for which data have been collected by Giger and Klüber for their project on SWISS MPS AND THEIR INTEREST GROUP AFFILIATION. Based on the Comparative Agenda Project (CAP) coding scheme.	CWA

Note on variables: The detailed description of variables, their ranges or categories are only provided for original variables presented in this codebook. For variables from secondary sources that are mentioned in the above table, readers are referred to the original codebooks and studies. The variables in question all start with a capitalised indicator: GAVA, NCCR, and GK.

org_id unique organisation identifier: unique identification code for organisations consisting of the country abbreviation (variable “country_abb”), and a 12 digit combination of letters, numbers, and special characters. The 12 digits code is created in Excel with the VBA hash function BASE64SHA1.

org_name organisation name Name(s) of the organisation as used over time in different languages. The string captures exclusively names that refer to exactly the same organisation (Organisations that split or merged are not added to the same organisation). Alternative names are separated by slashes. If available, the place where an organisation is headquartered is given at the end separated by a comma: [name₁]/[name₂]/[name_n], [placeofheadquarters].

country_abb country abbreviation of the parliament: abbreviated name of the country the parliament/district belongs to. The used abbreviations follow the ISO “ALPHA-2 code. (See: http://www.nationsonline.org/oneworld/country_code_list.htm)

legal_form legal form: the legal form of organisations as indicated in the data source. The Register of Interest Ties prepared by the Swiss Parliamentary Services for instance distinguishes between 15 different

types of legal forms. The current list of abbreviations and their meaning can be looked up online: <https://www.parlament.ch/centers/documents/de/interessen-nr.pdf>

polactive_start_date First known political activity: Captures the time when an organisation was first politically active e.g. had first ties to a parliamentarian. The date consists of the two-digit day, the month (first three letters, small, of the English name of the month) and the four-digit year. If only the month is known the the days are dropped and of only the years are know the day and month letters are dropped.

polactive_end_date Last known political activity: Captures the time when an organisation was last politically active e.g. had last ties to a parliamentarian. The date consists of the two-digit day, the month (first three letters, small, of the English name of the month) and the four-digit year. If only the month is known the the days are dropped and of only the years are know the day and month letters are dropped.

16 The Quota Module (QUOT)

This is a PCC data module on (gender) quota information (at the level of the national parties, **QUOT**). The quota data is appended to and compatible with the FACT data frame, however the data is strictly speaking not part of the PCC data. Responsible for the module and the data contained are: EF and TTZ.

Table 24: Summary of Quota Variables (QUOT)

NAME	TYPE	VALUES/EXAMPLE ³²	SHORT DESCRIPTION	EFFORT
faction_id	String	DE_NT_2009- DE_CDU_NT	faction identifier: combination of the parliament_id, and the party id (see 10).	COMP
quota_bin	Binary	0=no gender quota 1=gender quota	quota dummy: indicates whether the party has any ('soft' or 'hard') gender quota or not	CWA
quota_percentage	Numeric	0:100	quota value: mentions the percentage (of candidates or posts) that the quota specifies to be female at least	CWA
quota_zipper	Binary	0=no zipper quota 1=zipper quota	zipper quota dummy: indicates whether the party reserves alternating seats to different sexes	CWA
quota_soft	Binary	0='hard' quota 1='soft' quota	soft gender quota: shows if a gender quota is 'soft'/informal/non binding in wording or not	CWA
quota_execution	Binary	0=quota not implemented 1=quota implemented	quota implementation: indicates whether the party actually implemented the quota or not	CWA

faction_id faction identifier: combination of the parliament_id. and the name of the party (**faction_id**) (see 10). Identifies the party-parliament dyad for which the quota of the national party applies. In case the quota was introduced during a parliamentary term, the quota information is recorded for the next parliamentary term/election.

quota_bin quota dummy: Dummy variable indicating if a party/fraction in or for a certain parliament has any gender quota ('soft' or 'hard') (1) or not (0). Gender quotas are defined as any ratio/share of females to males (bigger than zero) with regard to electoral candidates, list postions, or elected MPs (party offices and delegates for party conventions are not considered) mentioned in an official party document (party statute, party convention decision, minutes of party convention) as officially decided by the party.

quota_females quota value: Numeric variable that mentions the highest percentage of females (candidates or positions) that the quota prescribes explicitly. When the quota text prescribes "at least X percent", X is the recorded value. The variable is zero if the party has no gender quota (see variable **quota_bin**). When the wording 'equal' is used this is coded as a 50% quota.

quota_zipper zipper quota dummy: Binary variable that indicates if (1) the respective party has a gender quota that reserves alternating positions on electoral lists for male and female candidates (zipper system). A zipper system can also leave certain positions open to both sexes (e.g. SPD in Germany, prescribes the electoral list positions to be alternating in sexes but leaves every fifth position open for either sex).

quota_soft soft gender quota: Dummy variable that shows if a gender quota is ‘soft’/non binding in wording (1) or not (0). For ‘soft’ quotas the party statutes say “ideally there should be...” or “there should be...” a certain share of females (or a similar expression).

quota_execution quota implementation: The dummy variable indicates whether, according to our information, the gender quota was implemented by the party in the respective legislative election/period (1) or not (0).

17 Appendices

17.1 Alternative IDs: ALID

ALID Different data sources offer different challenges and possibilities for the creation of IDs. To create an ID based on country, last name, first name, and year of birth might not be always be feasible. At the same time, IDs of existing data sets (e.g. the Manow IDs) might need to be transposed to the PCC IDs by some means. If the creation of the PCC is not possible in one step, the use of auxiliary IDs makes sense.

The following table offers an overview of alternative IDs and the alternative values they use. For all alternative IDs, a table of correspondence will then be created to easily allow for them to be transposed to the PCC IDs. Please note that these IDs are not a data-frame in the classic sense, they will just be generated on basis of the information available.

Table 25: Alternative IDs

Alternative Value(s)	ID Structure	Examples
Region	Country_LastName_FirstName_Region	CH_Allenspach_Heinz_ZH, DE_Doerflinger_Thomas_BW
Party	Country_LastName_FirstName_Party	CH_Allenspach_Heinz_FDP—PRD, DE_Doerflinger_Thomas_CDU, NL_Elias_Ton_VVD
Year of Parliament Entry	Country_LastName_FirstName_EntryYear	CH_Allenspach_Heinz_E1979, DE_Doerflinger_Thomas_E1998, NL_Elias_Ton_E2008

17.2 Regional Abbreviations Used

17.2.1 Regional Abbreviations Used for CH

- AG = Aargau
- AI = Appenzell Innerrhoden
- AR = Appenzell Ausserrhoden
- BS = Basel-Stadt
- BL = Basel-Landschaft
- BE = Bern
- FR = Fribourg
- GE = Geneva
- GL = Glarus
- GR = Graubünden
- JU = Jura
- LU = Lucerne
- NE = Neuchâtel
- NW = Nidwalden
- OW = Obwalden
- SH = Schaffhausen
- SZ = Schwyz
- SO = Solothurn
- SG = St. Gallen
- TG = Thurgau
- TI = Ticino
- UR = Uri
- VS = Valais

- VD = Vaud
- ZG = Zug

- ZH = Zürich

17.2.2 Regional Abbreviations Used for DE

- BW = Baden-Württemberg
- BY = Bavaria
- BE = Berlin
- BB = Brandenburg
- HB = Bremen
- HH = Hamburg
- HE = Hesse
- NI = Lower Saxony

- MV = Mecklenburg-Vorpommern
- NW = North Rhine-Westphalia
- RP = Rhineland-Palatinate
- SL = Saarland
- SN = Saxony
- ST = Saxony-Anhalt
- SH = Schleswig-Holstein
- TH = Thuringia

17.2.3 Regional Abbreviations Used for NL

- DR = Drenthe
- FL = Flevoland
- FR = Friesland
- GD = Gelderland
- GR = Groningen
- LB = Limburg
- NB = Noord-Brabant

- NH = Noord-Holland
- OV = Overijssel
- UT = Utrecht
- ZH = Zuid-Holland
- ZL = Zeeland
- BO = Openbare lichamen Bonaire, Sint Eustatius en Saba